FIGURE 1

GTTACTGCACCCCAAAACAGGTCTGGCCACGGCCATGAGCATGCTGAGCCATC**ATG**CCCACCGTGGATGACATTCTG
GAGCAGGTTGGGGAGTCTGGCTGGTTCCAGAAGCAAGCCTTCCTCATCTTATGCCTGCTGTCGGCTGCCTTTGCGCC
CATCTGTGTGGGCATCGTCTTCCTGGGTTTCACACCTGACCACCACTGCCAGAGTCCTGGGGTGGCTGAGCTGAGCC
AGCGCTGTGGCCTGGAGCCCTGCGGAGGAGCTGAACTATACAGTGCCAGGCCTGGGGCCCGCGGGCGAGGCCTTCCTT
GGCCAGTGCAGGCGCTATGAAGTGGACTGGAACCAGAGCGCCCTCAGCTGTGTAGACCCCCTGGCTAGCCTGGCCAC
CAACAGGAGCCACCTGCCGCTGGGTCCCTGCCAGGATGGCTGGGTGTATGACACGCCCGGCTCTTCCATCGTCACTG
AGTTCAACCTGGTGTGTGCTGACTCCTGGAAGCTGGACCTCTTTCAGTCCTGTTTGAATGCGGGCTTCTTGTTTGGC
TCTCTCGGTGTTGGCTACTTTGCAGACAGGTTTGGCCGTAAGCTGTGTCTCCTGGGAACTGTGCTGGTCAACGCGGT
GTCGGGCGTGCTCATGGCCTTCTCGCCCAACTACATGTCCATGCTGCTCTTCCGCCTGCTGCAGGGCCTGGTCAGCA
AGGGCAACTGGATGGCTGGCTACACCCTAATCACAGAATTTGTTGGCTCGGGCTCCAGAAGAACGGTGGCGATCATG
TACCAGATGGCCTTCACGGTGGGGCTGGTGGCGCTTACCGGGCTGGCCTACGCCCTGCCTCACTGGCGCTGGCTGCA
GCTGGCAGTCTCCCTGCCCACCTTCCTCTTCCTGCTCTACTACTGGTGTGTGGCCGGAGTCCCCTCGGTGGCTGTTAT
CACAAAAAAGAAACACTGAAGCAATAAAGATAATGGACCACATCGCTCAAAAGAATGGGAAGTTGCCTCCTGCTGAT
TTAAAGATGCTTTCCCTCGAAGAGGATGTCACCGAAAAGCTGAGCCCTTCATTTGCAGACCTGTTCCGCACGCCGCG
CCTGAGGAAGCGCCACCTTCATCCTGATGTACCTGTGGTTCACGGACTCTGTGCTCTATCAGGGCTCATCCTGCACA
TGGGCGCCACCAGCGGGAACCTCTACCTGGATTTCCTTTACTCCGCTCTGGTCGAAATCCCGGGGGCCTTCATAGCC
CTCATCACCATTGACCGCGTGGGCGCATCTACCCCATGGCCATGTCAAATTTGTTGGCGGGGGCAGCCTGCCTCGT
CATGATTTTTATCTCACCTGACCTGCACTGGTTAAACATCATAATCATGTGTGTTGGCCGAATGGGAATCACCATTG
CAATACAAATGATCTGCCTGGTGAATGCTGAGCTGTACCCCACATTCGTCAGGAACCTCGGAGTGATGGTGTGTTCC
TCCCTGTGTGACATAGGTGGGATAATCACCCCCTTCATAGTCTTCAGGCTGAGGGAGGTCTGGCAAGCCTTGCCCCT
CATTTTGTTTGCGGTGTTGGGCCTGCTTGCCGCGGGAGTGACGCTACTTCTTCCAGAGACCAAGGGGGTCGCTTTGC
CAGAGACCATGAAGGACGCCGAGAACCTTGGGAGAAAAGCAAAGCCCAAAGAAAACACGATTTACCTTAAGGTCCAA
ACCTCAGAACCCTCGGGCACC**TGA**GAGAGATGTTTTGCGGCGATGTCGTGTTGGAGGGATGAAGATGGAGTTATCCT
CTGCAGAAATTCCTAGACGCCTTCACTTCTCTGTATTCTTCCTCATACTTGCCTACCCCCAAATTAATATCAGTCCT
AAAGAAAAAAAAAAAAAAAAAAA

FIGURE 2

```
AATTAGCATAACCCTTCCTCAGGAAGAGTGAGATTTTATATTTGACAATAAAGTGTTAGACTCCATTTCTAAATACC
AGACTTCAAAAGATAAGGTTCAAAAGTGTTATAAGAAGATATTCCTTTTTTGTCCTAGAGAACTTATTTTCCTGTG
AAAATGCCTACCACAAAGAAGACATTGATGTTCTTATCAAGCTTTTTCACCAGCCTTGGGTCCTTCATTGTAATTTT
GCTCTATTCTTGGGACACAAGCATGGATCACCAGTACAATTGCTGTTAGAGACTCTGCTTCAAATGGGAGCATTTTC
ATCACTTACGGACTTTTTCGTGGGGAGAGTAGTGAAGAATTGAGTCACGGACTTGCAGAACCAAAGAAAAAGTTTGC
AGTTTTAGAGATACTGAATAATTCTTCCCAAAAAACTCTGCATTCGGTGACTATCCTGTTCCTGGTCCTGAGTTTGA
TCACGTCGCTGCTGAGCTCTGGGTTTACCTTCTACAAACAGCATCAGCAACCCTTACCAGACATTCCTGGGGCCCGA
CGGGGGTGTACACCTGGGAACGGGCTTCGGNGCATNCTTCGTTTTTGGTGACCATGATACTGGTTTGNGGGNAACAC
GCAGGTCCCACCCAACTCTTCCGAANAAGTTGTTTCCAAAATGCTTTACCCCCGGNAACCANCCAGTAAAGGAACGA
CCCCACAGGTACCGGATACTCGTGCTGGCTCATACTGG
```

FIGURE 3

GATCAGTGTGTGAGGGAACTGCCATC**ATG**AGGTCTGACAAGTCAGCTTTGGTATTTCTGCTCCTGCAGCTCTTCTGT
GTTGGCTGTGGATTCTGTGGGAAAGTCCTGGTGTGGCCCTGTGACATGAGCCATTGGCTTAATGTCAAGGTCATTCT
AGAAGAGCTCATAGTGAGGAGGCCATGAGGTAACAGTATTGACTCACTCAAAGCCTTCGTTAATTGACTACAGGAAGC
CTTCTGCATTGAAATTTGAGGTGGTCCATATGCCACAGGACAGAACAGAAGAAAATGAAATATTTGTTGACCTAGCT
CTGAATGTCTTGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAAATGATTTTTTTGTTGAAATAAGAGGAAC
TTTAAAAATGATGTGTGAAGACGTTTATCTACAATCAGACGCTTATGAAGAAGCTACAGGAAACCAACTACGATGTAA
TGCTTATAGACCCTGTGATTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGA
ATTTCTGTAGGAGGCAATATGGAGCGAAGCTGTGGGAAACTTCCAGCTCCACTTTCCTATGTACCTGTGCCTATGAC
AGGACTAACAGACAGAATGACCTTTCTGGAAAGAGTAAAAAATTCAATGCTTTCAGTTTTGTTCCACTTCTGGATTC
AGGATTACGACTATCATTTTTGGGAAGAGTTTTATAGTAAGGCATTAGGAAGGCCCACTACATTATGTGAGACTGTG
GGAAAAGCTGAGATATGGCTAATACGAACATATTGGGATTTTGAATTTCCTCAACCATACCAACCTAACTTTGAGTT
TGTTGGAGGATTGCACTGTAAACCTGCCAAAGCTTTGCCTAAGGAAATGGAAAATTTTGTCCAGAGTTCAGGGGAAG
ATGGTATTGTGGTGTTTTCTCTGGGGTCACTGTTTCAAAATGTTACAGAAGAAAAGGCTAATATCATTGCTTCAGCC
CTTGCCCAGATCCCACAGAAGGTGTTATGGAGGTACAAAGGAAAAAAACCATCCACATTAGGAGCCAATACTCGGCT
GTATGATTGGATACCCCAGAATGATCTTCTTGGTCATCCCAAAACCAAAGCTTTTATCACTCATGGTGGAATGAATG
GGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGGAGTTCCCATATTTGGTGATCAGCTTGATAACACATAGCT
CACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAAACTATGACAAGCGAAGATTTACTGAGGGCTTTGAG
AACAGTCATTACCGATTCCTCCTTATAAAGAGAATGCTATGAGATTATCAAGAATTCACCATGATCAACCTGTAAAGC
CCCTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCATGCGCCACAAAGGAGCCAAGCACCTGCGATCAGCTGCCCAT
GACCTCACCTGGTTCCAGCACTACTCTATAGATGTGATTGGGTTCCTGCTGACCTGTGTGGCAACTGCTATATTCTT
GTTCACAAAATGTTTTTTATTTTCCTGTCAAAAATTTAATAAAACTAGAAAGATAGAAAAGAGGGAA**TAG**ATCTTTC
CAAATTCAAGAAAGACCTGATGGGGTAATCCTGTTAATTCCAGCCACATAGAATTTGGTGAAAACCTTGCTATTTTC
ATATTATCTATTCTGTTATTTTATCTTAGCTATATAGCCTAGAATTCCATGATCATGAGGTTGTGAGTATATCTCAT
TCTTTCGTTGTATTTTCCTAGGTGTCTTTACTCTCTTCTCTCACTTTGTGACACAAGGACATGAATACATCTAAATT
TTCCTATTTCTGATATGACTGTTTTGATGATGTCATTACTTCTATAACCTTAAGTGATAGGGTGACATGCAATATGA
TTATTCCTGGTGTGCGCCCAAACACATGGATATAAAGAGGTAAAAAACTTAAAATTCACAAAATTCAGTAAACCACA
CAAATCAGGTAAGTGTTCTATGAGATTAGCTGGCTATGAGAAACATAATGATGTTTCTTTTTCAATTTAAATAAGCC
TTTCTACATAGCCAGCATCAGTGATCTCAGAAAATAAATTGCTAATAATGATGACATGGCATTATGCTTAGAAAAGT
TTGCTGTATTTCCATAGACCTCATCTAGATGTCATGGCCTACATTTCTGCCATCACTCAACCAATACTTTTTTCTGT
TTTCTTGATGATAAAAAGACCTTTCTCATGATTGCCATCAAATAACAAAAGAAACTATTTTTTTTTCTCACATAGAGA
ACATGTCAGTAAGATATTCAAGGTGAACAGATATTTTTGGGATTAGTAACTATTTGAAATATGTGGTGATAATTACT
GAGTTTATAAAATTTATTTGATAGTACACTTAAAGAAGATTTATATGTTTATTCTTTAAAAATGATGAATACTCATA
ATTCTTATCTCTATAATCAAAAGTATAATTTACTGTAGAAAAAATAAAGAGATGCTTGTTTTTCAATTTAAATAAGCC
AACTGCTTTTCAGTCTCAATCTTTGAGAATTGTAAATTCATCAAATAATTGCTTACATAGTAAAAATTTAAGGTATT
AGAAAACCTGCATAACAAATAGTATTATATATTAAATATTTTGATATGTAAAGCTCTACACAAAGCTAAAATATAGTG
TAATAATGTTTACACTAGTAAGCAAATATGTTAATCTTCTCATTTTTTTACTGTCATATAATCTTAGTGATATGCCT
ATTAATAGTTTTAAATAAATAAATTGGCTTATCTGGCTTTTTGAAAATTTTGAAATTCTTACAGATGTTGATTAGGT
ATATCTACAAATTAATTTCAATTTTAAAATGATGATATAAAAATAAATATAAGTATTTTTCTTGTGTATGTATACAA
TAAATATAAATAAAATTGTTTACTGTTTTGAAAGTTTCTTAAGTTTTA

FIGURE 4

GAGGATGCTGCTGGGGAAGATGTGCGCGCCGAGGTACCCGTCCGGGGGCGCCTGGGACCCTCACCCAGGCCAGGCTT
CGGGGAGAGCATGTATCTGCTCTCGGACAAGGCCACCTCGCCGCTCTCGCTGGATGCTGGCCTCGGGCAGGCCCCCT
GGAGCGACCTGCTTCTTTGGGCACTGTTGCTGAACAGGGCACAGATGGCCATGTACTTCTGGGAGATGGGTTCCAAT
GCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGGCACGCGTGGAGCCTGACGCTGAGGAGGCAGC
ACGGAGGAAAGACCTGGCGTTCAAGTTTGAGGGGATGGGCGTTGACCTCTTTGGCGAGTGCTATCGCAGCAGTGAGG
TGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCCCGCTCTGGGGGGGATGCCACTTGCCTCCAGCTGGCC**ATG**CAAGC
TGACGCCCGTGCCTTCTTTGCCCAGGATGGGGTACAGTCTCTGCTGACACAGAAGTGGTGGGGAGATATGGCCAGCA
CTACACCCATCTGGGCCCTGGTTCTCGCCTTCTTTTGCCCTCCACTCATCTACACCCGCCTCATCACCTTCAGGAAA
TCAGAAGAGGAGCCCACACGGGAGGAGCTAGAGTTTGACATGGATAGTGTCATTAATGGGGAAGGGCCTGTCGGGAC
GGCGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTCCGGGTTGCTGCGGGGGGCCGCTGCG
GGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGCGCCGGTGACCATCTTCATGGGCAACGTGGTCAGC
TACCTGCTGTTCCTGCTGCTTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGCCGGCGCCGCCCGGCTCCCTGGAGCT
GCTGCTCTATTTCTGGGCTTTCACGCTGCTGTGCGAGGAACTGCGGCCAGGGCCTGAGCGGAGGCGGGGGCAGCCTCG
CCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCTACCTCGCCGACAGCTGGAACCAG
TGCGACCTAGTGGCTCTCACCTGCTTCCTCCTGGGCGTGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCG
CACTGTCCTCTGCATCGACTTCATGGTTTTCACGGTGCGGCTGCTTCACATCTTCACGGTCAACAAACAGCTGGGGC
CCAAGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTCTTCCTCGGCGTGTGGCTGGTAGCC
TATGGCGTGGCCACGGAGGGGCTCCTGAGGCCACGGGACAGTGACTTCCCAAGTATCCTGCGCCGCGTCTTCTACCG
TCCCTACCTGCAGATCTTCGGGCAGATTCCCCAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGT
CGGAGCCCGGCTTCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGCTGGTG
GTGCTGCTCCTCGTCATCTTCCTGCTCGTGGCCAACATCCTGCTGGTCAACTTGCTCATTGCCATGTTCAGTTACAC
ATTCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGAAGGCGCAGGTTACCGCCTCATCCGGGAATTCCACTCTCG
GCCCGCGCTGGCCCCGCCCTTTATCGTCATCTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGACCCCGGA
GCCCCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGCGGAAGTGCTAACGT
GGGAATCGGTGCA**TAA**GGAGAACTTTCTGCTGGCACGCGCTAGGGACAAGCGGGAGAGCGACTCCGAGCGTCTGAAG
CGCACGTCCCAGAAGGTGGACTTGGCACTGAAACAGCTGGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCT
GGAGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGCTGCCCC
CAGGTGGGCCGCCACCCCCTGACCTGCCTGGGTCCAAAGACTGAGCCCTGCTGGCCGGACTTCAAGGAGAAGCCCCCA
CAGGGGATTTTGCTCCTAGAGTAAGGCTCATCTGGGCCTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGC
CCCATGTCCATCTGGGCCACTGTCAGGACCACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGGCTCCTC
CCAGAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTATCCATCTGGAGGCTGCAGGGTCCTTG
GGGTAACAGGGACCACAGACCCCTCACCACTCACAGATTCCTCACACTGGGGAAATAAAGCCATTTCAGAGG

FIGURE 5

TNCCCGCAGAACCAGGAAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCCGGCTAGAAAAACTCTGTCGG
TACCAACCCCAGAGCGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGAGAGGTGCAGGACTCAGACTTCACCAG
CCCACTCGGTCCCAGCCTTGTACGCAAAGAGACGCCAAGGACGCGCTCTCCCGCGTCCAGGCAGCCCCAGCTTGCTG
GCTTGCCTGCCCGCCTGCGTGCAGCACTCGGCCGGCGTGCAGC**ATG**ACCCTGTGGAACGGCGTACTGCCTTTTTACC
CCCAGCCCCGGCATGCCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTAGCAGCAAGC
TTCCTGCTCATCTTGCCGGGGATCCGTGGCCACTCGCGCTGGTTTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTCAT
AGGCGCAGAAATTGTGGCTGTGCACTTCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTACAAAG
CCTTCAGCGCAGCGCGCGTTACAGCCCGTGTCGGTCTGCTCGTGGGCCTGGAGGGCATTAATATTACACTCACAGGG
ACCCCAGTGCATCAGCTGAACGAGACCATTGACTACAACGAGCAGTTCACCTGGCGTCTGAAAGAGAATTACGCCGC
GGAGTACGCGAACGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTCTACCTGGCGGAGAAGTTCACACCGAGTAGCC
CTTGCGGCCTGTACCACCAGTACCACCTGGCGGGACACTACGCCTCGGCCACGCTATGGGTGGCGTTCTGCTTCTGG
CTCCTCTCCAACGTGCTGCTCTCCACGCCGGCCCCGCTCTACGGAGGCCTGGCACTGCTGACCACCGGAGCCTTCGC
GCTCTTCGGGGTCTTCGCCTTGGCCTCCATCTCTAGCGTGCCGCTCTGCCCGCTCCGCCTAGGCTCCTCCGCGCTCA
CCACTCAGTACGGCGCCGCCTTCTGGGTCACGCTGGCAACCGGCGTCCTGTGCCTCTTCCTCGGAGGGGCCGTGGTG
AGTCTCCAGTATGTTCGGCCCAGCGCTCTTCGCACCCTTCTGGACCAAAGCGCCAAGGACTGCAGCCAGGAGAGAGG
GGGCTCACCTCTTATCCTCGGCGACCCACTGCACAAGCAGGCCGCTCTCCCAGACTTAAAATGTATCACCACTAACC
TG**TGA**GGGGGACCCAATCTGGACTCCTTCCCCGCCTTGGGACATCGCAGGCCGGGAAGCAGTGCCCGCCAGGCCTGG
GCCAGGAGAGCTCCAGGAAGGGCACTGAGCGCTGCTGGCGCGAGGCCTCGGACATCCGCAGGCACCAGGGAAAGTCT
CCTGGGGCGATCTGTAAATAAACCTTTTTTTCTTTTGTTTTTTAAAAAAAAAATAAAAGTCGACC

FIGURE 6

```
CTGCAGGCTTCAGGAGGGGACACAAGCATGGAGCGGCTTTGGGGTCTATTCCAGAGAGCGCAACAACTGTCCCCAAG
ATCCTCTCAGACCGTCTACCAGCGTGTGGAAGGCCCCCGGAAAGGGCACCTGGAGGAGGAAGAGGAAGACGGGGAGG
AGGGGGCGGAGACATTGGCCCACTTCTGCCCCATGGAGCTGAGGGGCCCTGAGCCCCTGGGCTCTAGACCCAGGCAG
CCAAACCTCATTCCCTGGGCGGCAGCAGGACGAGGGCTGCCCCCTACCTGGTCCTGACGGCCCTGCTGATCTTCAC
TGGGGCCTTCCTACTGGGCTACGTCGCCTTCCGAGGGTCCTGCCAGGCGTGCGGAGACTCTGTGTTGGTGGTCAGTG
AGGATGTCAACTATGAGCCTGACCTGGATTTCCACCAGGGCAGACTCTACTGGAGCGACCTCCAGGCCATGTTCCTG
CAGTTCCTGGGGGAGGGGCGCCTGGAGGACACCATCAGGCAAACCAGCCTTCGGGAACGGGTGGCAGGCTCGGCCGG
GATGGCCGCTCTGACTCAGGACATTCGCGCGGCGCTCTCCCGCCAGAAGCTGGACCACGTGTGGACCGACACGCACT
ACGTGGGGCTGCAATTCCCGGATCCGGCTCACCCCAACACCCTGCACTGGGTCGATGAGGCCGGGAAGGTCGGAGAG
CAGCTGCCCGCTGGAGGACCCTGACGTCTACTGCCCCTACAGCGCCATCGGCAACGTCACGGGAGAGCTGGTGTACGC
CCACTACGGGCGGCCCGAAGACCTGCAGGACCTGCGGGGCCAGGGGCGTGGATCCAGTGGGCCGCCTGCTGCTGGTGC
GCGTGGGGGTGATCAGCTTCGCCCAGGAAGGTGACCAATGCTCAGGACTTCGGGGGCTCAAGGAGTGCTCATATACCCA
GAGCCAGCGGACTTCTCCCAGGACCCACCCAAGCCAAGCCTGTCCAGCCAGCAGGCAGTGTATGGACATGTGCACCT
GGGAACTGGAGACCCCTACACACCTGGCTTCCCTTCCTTCAATCAAACCCAGTTCCCTCCAGTTGCATCATCAGGCC
TTCCCAGCATCCCAGCCCAGCCCATCAGTGCAGACATTGCCTCCCGCCTGCTGAGGAAGCTCAAAGGCCCTGTGGCC
CCCCAAGAATGGCAGGGGAGCCTCCTAGGCTCCCCTTATCACCTGGGCCCCGGGCCACGACTGCGGCTAGTGGTCAA
CAATCACAGGACCTCCACCCCATCAACAACATCTTCGGCTGCATCGGAAGGCCGCTCAGAGCCAGATCACTACGTTG
TCATCGGGGCCCAGAGGGATGCATGGGGCCCAGGAGCAGCTAAATCCGCTGTGGGGACGGCTATACTCCTGGAGCTG
GTGCGGACCTTTTCCTCCATGGTGAGCAACGGCTTCCGGCCCCGCAGAAGTCTCCTCTTCATCAGCTGGGACGGTGG
TGACTTTGGAAGCGTGGGCTCCACGGAGTGGCTAGAAGGCTACCTCAGCGTGCTGCACCTCAAAGCCGTAGTGTACG
TGAGCCTGGACAACGCAGTGCTGGGGGATGACAAGTTTCATGCCAAGACCAGCCCCCTTCTGACAAGTCTCATTGAG
AGTGTCCTGAAGCAGGTGGATTCTCCCAACCACAGTGGGCAGACTCTCTATGAACAGGTGGTGTTCACCAATCCCAG
CTGGGATGCTGAGGTGATCCGGCCCCTACCCATGGACAGCAGTGCCTATTCCTTCACGGCCTTTGTGGGAGTCCCTG
CCGTCGAGTTCTCCTTTATGGAGGACGACCAGGCCTACCCATTCCTGCACACAAAGGAGGACACTTATGAGAACCTG
CATAAGGTGCTGCAAGGCCGCCTGCCCGCCGTGGCCCAGGCCGTGGCCCAGCTCGCAGGGCAGCTCCTCATCCGGCT
CAGCCACGATCGCCTGCTGCCCCTCGACTTCGGCCGCTACGGGGACGTCGTCCTCAGGCACATCGGGAACCTCAACG
AGTTCTCTGGGGACCTCAAGGCCCGCGGGCTGACCCTGCAGTGGGTGTACTCGGCGCGGGGGACTACATCCGGGCG
GCGGAAAAGCTGCGGCAGGAGATCTACAGCTCGGAGGAGAGAGACGAGCGACTGACACGCATGTACAACGTGCGCAT
AATGCGGGTGGAGTTCTACTTCCTTTCCCAGTACGTGTCGCAGCCGACTCCCCGTTCCGCCACATCTTCATGGGCC
GTGGAGACCACACGCTGGGCGCCCTGCTGGACCACCTGCGGCTGCTGCGCTCCAACAGCTCCGGGACCCCCGGGGCC
ACCTCCTCCACTGGCTTCCAGGAGAGCCGTTTCCGGCGTCAGCTAGCCCTGCTCACCTGGACGCTGCAAGGGGCAGC
CAATGCGCTTAGCGGGGATGTCTGGAACATTGATAACAACTTCTGAGGCCCTGGGGATCCTCACATCCCCGTCCCCC
AGTCAAGAGCTCCTCTGCTCCTCGCTTGAATGATTCAGGGTCAGGGAGGTGGCTCAGAGTCCACCTCTCATTGCTGA
TCAATTTCTCATTACCCCTACACATCTCTCCACGGAGCCCAGACCCCAGCACAGATATCCACACACCCCAGCCCTGC
AGTGTAGCTGACCCTAATGTGACGGTCATACTGTCGGTTAATCAGAGAGTAGCATCCCTTCAATCACAGCCCCTTCC
CCTTTCTGGGGTCCTCCATACCTAGAGACCACTCTGGGAGGTTTGCTAGGCCCTGGGACCTGGCCAGCTCTGTTAGT
GGGAGAGATCGCTGGCACCATAGCCTTATGGCCAACAGGTGGTCTGTGGTGAAAGGGGCGTGGAGTTTCAATATCAA
TAAACCACCTGATATCAATAAGCCAAAA
```

FIGURE 7

GCTGGAGCATCCCGCTCTGGTGCCGCTGCAGCCGGCAGAG**ATG**GTTGAGCTCATGTTCCCGCTGTTGCTCCTCCTTC
TGCCCTTCCTTCTGTATATGGCTGCGCCCCAAATCAGGAAAATGCTGTCCAGTGGGGTGTGTACATCAACTGTTCAG
CTTCCTGGGAAAGTAGTTGTGGTCACAGGAGCTAATACAGGTATCGGGAAGGAGACAGCCAAAGAGCTGGCTCAGAG
AGGAGCTCGAGTATATTTAGCTTGCCGGGATGTGGAAAAGGGGGAATTGGTGGCCAAAGAGATCCAGACCACGACAG
GGAACCAGCAGGTGTTGGTGCGGAAACTGGACCTGTCTGATACTAAGTCTATTCGAGCTTTTGCTAAGGGCTTCTTA
GCTGAGGAAAAGCACCTCCACGTTTTGATCAACAATGCAGGAGTGATGATGTGTCCGTACTCGAAGCAGCAGATGG
CTTTGAGATGCACATAGGAGTCAACCACTTGGGTCACTTCCTCCTAACCCATCTGCTGCTAGAGAAACTAAAGGAAT
CAGCCCCATCAAGGATAGTAAATGTGTCTTCCCTCGCACATCACCTGGGAAGGATCCACTTCCATAACCTGCAGGGC
GAGAAATTCTACAATGCAGGCCTGGCCTACTGTCACAGCAAGCTAGCCAACATCCTCTTCACCCAGGAACTGGCCCG
GAGACTAAAAGGCTCTGGCGTTACGACGTATTCTGTACACCCTGGCACAGTCCAATCTGAACTGGTTCGGCACTCAT
CTTTCATGAGATGGATGTGGTGGCTTTTCTCCTTTTTCATCAAGACTCCTCAGCAGGGAGCCCAGACCAGCCTGCAC
TGTGCCTTAACAGAAGGTCTTGAGATTCTAAGTGGGAATCATTTCAGTGACTGTCATGTGGCATGGGTCTCTGCCCA
AGCTCGTAATGAGACTATAGCAAGGCGGCTGTGGGACGTCAGTTGTGACCTGCTGGGCCTCCCAATAGAC**TAA**CAGG
CAGTGCAGTTGGACCCAAGAGAAGACTGCAGCAGACTACACAGTACTTCTTGTCAAAATGATTCTCCTTCAAGGTTT
TCAAAACCTTTAGCACAAAGAGAGCAAAACCTTCCAGCCTTGCCTGCTTGGTGTCCAGTTAAAACTCAGTGTACTGC
CAGATTCGTCTAAATGTCTGTCATGTCCAGATTTACTTTGCTTCTGTTACTGCCAGAGTTACTAGAGATATCATAAT
AGGATAAGAAGACCCTCATATGACCTGCACAGCTCATTTTCCTTCTGAAAGAAACTACTACCTAGGAGAATCTAAGC
TATAGCAGGGATGATTTATGCAAATTTGAACTAGCTTCTTTGTTCACAATTCAGTTCCTCCCAACCAACCAGTCTTC
ACTTCAAGAGGGCCACACTGCAACCTCAGCTTAACATGAATAACAAAGACTGGCTCAGGAGCAGGGCTTGCCCAGGC
ATGGTGGATCACCGGAGTCAGTAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCACCTCTACTAAAAATTGTG
TATATCTTTGTGTGTCTTCCTGTTTATGTGTGCCAAGGGAGTATTTTCACAAAGTTCAAAACAGCCACAATAATCAG
AGATGGAGCAAACCAGTGCCATCCAGTCTTTATGCAAATGAAATGCTGCAAAGGGAAGCAGATTCTGTATATGTTGG
TAACTACCCACCAAGAGCACATGGGTAGCAGGGAAGAAGTAAAAAAAGAGAAGGAGAATACTGGAAGATAATGCACA
AAATGAAGGGACTAGTTAAGGATTAACTAGCCCTTTAAGGATTAACTAGTTAAGGATTAATAGCAAAAGACATTAAA
TATGCTAACATAGCTATGGAGGAATTGAGGGCAAGCACCCAGGACTGATGAGGTCTTAACAAAAACCAGTGTGGCAAA

FIGURE 8

```
GCTCGGCCAGGCTGCTGGTACCTGCGTCCGCCCGGCGAGCAGGACAGGCTGCTTTGGTTTGTGACCTCCAGGCAGGA
CGGCCATCCTCTCCAGAATGAAGATCTTCTTGCCAGTGCTGCTGGCTGCCCTTCTGGGTGTGGAGCGAGCCAGCTCG
CTGATGTGCTTCTCCTGCTTGAACCAGAAGAGCAATCTGTACTGCCTGAAGCCGACCATCTGCTCCGACCAGGACAA
CTACTGCGTGACTGTGTCTGCTAGTGCCGGCATTGGGAATCTCGTGACATTTGGCCACAGCCTGAGCAAGACCTGTT
CCCCGGCCTGCCCCATCCCAGAAGGCGTCAATGTTGGTGTGGCTTCCATGGGCATCAGCTGCTGCCAGAGCTTTCTG
TGCAATTTCAGTGCGGCCGATGGCGGGCTGCGGGCAAGCGTCACCCTGCTGGGTGCCGGGCTGCTGCTGAGCCTGCT
GCCGGCCCTGCTGCGGTTTGGCCCCTGACCGCCCAGACCCTGTCCCCCGATCCCCCAGCTCAGGAAGGAAAGCCCAG
CCCTTTCTGGATCCCACAGTGTATGGGAGCCCCTGACTCCTCACGTGCCTGATCTGTGCCCTTGGTCCCAGGTCAGG
CCCACCCCCTGCACCTCCACCTGCCCCAGCCCCTGCCTCTGCCCCAAGTGGGGCCAGCTGCCCTCACTTCTGGGGTG
GATGATGTGACCTTCCTTGGGGGACTGCGGAAGGGACGAGGGTTCCCTGGAGTCTTACGGTCCAACATCAGGACCAA
GTCCCATGGACATGCTGACAGGGTCCCCAGGGAGACCGTGTCAGTAGGGATGTGTGCCTGGCTGTGTACGTGGGTGT
GCAGTGCACGTGAGAGCACGTGGCGGCTTCTGGGGGCCATGTTTGGGGAGGGAGGTGTGCCAGCAGCCTGGAGAGCC
TCAGTCCCTGTAGCCCCCTGCCCTGGCACAGCTGCATGCACTTCAAGGGCAGCCTTTGGGGGTTGGGGTTTCTGCCA
CTTCCGGGTCTAGGCCCTGCCCCAAATCCAGCCAGTCCTGCCCCCAGCCCACCCCCACATTGGAGCCCTCCTGCTGCT
TTGGTGCCTCAAATAAATACAGATGTCCCCCAAAAAAAAAAAAAAAAAA
```

FIGURE 9

GAAGTTTCTCACTAGGGTCTTCTCTGGCCCAGCCTTTGACTGAAGCTGGTCTGGAGACAGGGGCATTAGAGAAGTGA
CTCATAGATGGCCTAAAGAAGCGGGGCCACTCAAGGACCCAGGACAGAGGGAAGAGGGCCAACCCAGCTGGACCACA
GGCAAACCCCATTGCCTTTGAGAGAAAGAAGAGGACCCGGTGAAAC**ATG**CTGCTGCTGAAGAAACACACGGAGGACA
TCAGCAGCGTCTACGAGATCCGCGAGAGGCTCGGCTCGGGTGCCTTCTCCGAGGTGGTGCTGGCCCAGGAGCGGGGC
TCCGCACACCTCGTGGCCCTCAAGTGCATCCCCAAGAAGGCCCTCCGGGGCAAGGAGGCCCTGGTGGAGAACGAGAT
CGCAGTGCTCCGTAGGATCAGTCACCCCAACATCGTCGCTCTGGAGGATGTGCCACGAGAGCCCTTCCCACCTCTACC
TGGCCATGGAACTGGTGACGGGTGGCGAGCTGTTTGACCGCATCATGGAGCGCGGCTCCTACACAGAGAAGGATGCC
AGCCATCTGGTGGGTCAGGTCCTTGGCGCCGTCTCCTACCTGCACAGCCTGGGGATCGTGCACCGGGACCTCAAGCC
CGAAAACCTCCTGTATGCCACGCCCTTTGAGGACTCGAAGATCATGGTCTCTGACTTTGGACTCTCCAAAATCCAGG
CTGGGAACATGCTAGGCACCGCCTGTGGGACCCCTGGATATGTGGCCCCAGAGCTCTTGGAGCAGAAACCCTACGGG
AAGGCCGTAGATGTGTGGGCCCTGGGCGTCATCTCCTACATCCTGCTGTGTGGGTACCCCCCCTTCTACGACGAGAG
CGACCCTGAGCTCTTCAGCCAGATCCTGAGGGCCAGCTATGAGTTTGACTXTCCTTTCTGGGATGACATCTCAGAAT
CAGGCAAAGACTTTATTCGGCACCTTCTGGAGCGAGACCTTCAGAAGAGGTTCACCTGCCAACAGGCCTTGCGGGAC
CTTTGGATCTTTTGGGACACAGGCTTTGGCAGGGACATCTTAGGGTTTGTCAGTGAGCAGATCCGGAAGAACTTTGC
TTGGACACACTGGAAGCGAGCCTTCAATGCCACCTTGTTCCTGCGCCACATCCGGAAGCTGGGGCAGATCCCAGAGG
GCGAGGGGGCCTCTGAGCAGGGCATGGSCCGXCACAGCCACTXAGGCCTTCGTGCTGGCCAGCCCCCCAAGTGG**TGA**
TGCCCAGGXAGATGCCGAGGCCAAGTGGAXTGAXCCCCAGATTTXCTTXC

FIGURE 10

GGAACGAGTGGGAACGTAGCTGGTCGCAGAGGGCACCAGCGGCTGCAGGACTTCACCAAGGGACCCTGAGGCTCGTG
AGCAGGGACCCGCGGTGCGGGTTATGCTGGGGGCTCAGATCACCGTAGACAACTGGACACTCAGGACCACGCC**ATG**G
AGGAGCTGCAGGATGATTATGAAGACATGATGGAGGAGAATCTGGAGCAGGAGGAATATGAAGACCCAGACATCCCC
GAGTCCCAGATGGAGGAGCCGGCAGCTCACGACACCGAGGCAACAGCCACAGACTACCACACCACATCACACCCGGG
TACCCACGAGGTCTATGTGGAGCTGCAGGAGCTGGTGATGGACGAAAAGAACCAGGAGCTGAGATGGATGGAGGCGG
CGCGCTGGGTGCAACTGGAGGAGAACCTGGGGGAGAATGGGGGCCTGGGGGCCGCCCGCACCTCTCTCACCTCACCTTC
TGGAGCCTCCTAGAGCTGCGTAGAGTCTTCACCAAGGGTACTGTCCTCCTAGACCTGCAAGAGACCTCCCTGGCTGG
AGTGGCCAACCAACTGCTAGACAGGTTTATCTTTGAAGACCAGATCCGGCCTCAGGACCGAGAGGAGCTGCTCCGGG
CCCTGCTGCTTAAACACAGCCACGCTGGAGAGCTGGAGGCCCTGGGGGGTGTGAAGCCTGCAGTCCTGACACGCTCT
GGGGATCCTTCACAGCCTCTGCTCCCCCAACACTCCTCACTGGAGACACAGCTCTTCTGTGAGCAGGGAGATGGGGG
CACAGAAGGGCACTCACCATCTGGAATTCTGGAAAAGATTCCCCCGGATTCAGAGGCCACGTTGGTGCTAGTGGGCC
GCGCCGACTTCCTGGAGCAGCCGGTGCTGGGCTTCGTGAGGCTGCAGGAGGCAGCGGAGCTGGAGGCGGTGGAGCTG
CCGGTGCCTATACGCTTCCTCTTTGTGTTGCTGGGACCTGAGGCCCCCACATCGATTACACCCAGCTTGGCCGGGGC
TGCTGCCACCCTCATGTCAGAGAGGGTGGTTCCGCATAGATGCCTACATGGCTCAGAGCCGAGGGGAGCTGCTGCACT
CCCTAGAGGGCTTCCTGGACTGCAGCCTAGTGCTGCCTCCCACCGATGCCCCCTCCGAGCAGGCACTGCTCAGTCTG
GTGCCTGTGCAGAGGGAGCTACTTCGAAGGCGCTATCAGTCCAGCCCTGCCAAGCCAGACTCCAGCTTCTACAAGGG
CCTAGACTTAAATGGGGGCCCAGATGACCCTCTGCAGCAGACAGGCCAGCTCTTCGGGGGCCTGGTGCGTGATATCC
GGCGCCGCTACCCCTATTACCTGAGTGACATCACAGATGCATTCAGCCCCCAGGTCCTGGCTGCCGTCATCTTCATC
TACTTTGCTGCACTGTCACCCGCCATCACCTTCGGCGGCCTCCTGGGAGAAAAGACCCGGAACCAGATGGGAGTGTC
GGAGCTGCTGATCTCCACTGCAGTGCAGGGCATTCTCTTCGCCCTGCTGGGGGCTCAGCCCCTGCTTGTGGTCGGCT
TCTCAGGACCCCTGCTGGTGTTTGAGGAAGCCTTCTTCTCGTTCTGCGAGACCAACGGTCTAGAGTACATCGTGGGC
CGCGTGTGGATCGGCTTCTGGCTCATCCTGCTGGTGGTGTTGGTGGTGGCCTTCGAGGGTAGCTTCCTGGTCCGCTT
CATCTCCCGCTATACCCAGGAGATCTTCTCCTTCCTCATTTCCCTCATCTTCATCTATGAGACTTTCTCCAAGCTGA
TCAAGATCTTCCAGGACCACCCACTACAGAAGACTTATAACTACAACGTGTTGATGGTGCCCAAACCTCAGGGCCCC
CTGCCCAACACAGCCCTCCTCTCCCTTGTGCTCATGGCCGGTACCTTCTTCTTTGCCATGATGCTGCGCAAGTTCAA
GAACAGCTCCTATTTCCCTGGCAAGCTGCGTCGGGTCATCGGGGACTTCGGGGGTCCCCATCTCCATCCTGATCATGG
TCCTGGTGGATTTCTTCATTCAGGATACCTACACCCCAGAAACTCTCGGTGCCTGATGGCTTCAAGGTGTCCAACTCC
TCAGCCCGGGGCTGGGTCATCCACCCACTGGGGCTTGCGTTCCGAGTTTCCCATCTGGATGATGTTTGCCTCCGCCCT
GCCTGCTCTGCTGGTCTTCATCCTCATATTCCTGGAGTCTCAGATCACCACGCTGATTGTCAGCAAACCTGAGCGCA
AGATGGTCAAGGGCTCCGGCTTCCACCTGGACCTGCTGCTGGTAGTAGGCATGGGTGGGGTGGCCGGCCCTCTTTGGG
ATGCCCTGGCTCAGTGCCACCACCGTGCGTTCCGTCACCCATGCCAACGCCCTCACTGTCATGGGCAAAGCCAGCAC
CCCAGGGGCTGCAGCCCAGATCCAGGAGGTCAAAGAGCAGCGGATCAGTGGACTCAGTGGACTCCTGGTCGCTGTGCTTGTGGGCC
TGTCCATCCTCATGGAGCCCATCCTGTCCCGCATCCCCCTGGCTGTACTGTTTGGCATCTTCCTCTACATGGGGGTC
ACGTCGCTCAGCGGCATCAGCTCTTTGACCGCATCTTGCTTCTGTTCAAGCCACCCAAGTATCACCCAAGATGTGCC
CTACGTCAAGCGGGTGAAGACCTGGCGCATGCACTTATTCACGGGCATCCAGATCATCTGCCTGGCAGTGCTGTGGG
TGGTGAAGTCCACGCCGGCCTCCCTGGCCCTGCCCTTCGTCCTCATCCTCACTGTGCCGCTGCGGCGCGCTCCTGCTG
CCGCTCATCTTCAGGAACGTGGAGCTTCAGTGTCTGGATGCTGATGATGCCAAGGCAACCTTTGATGAGGAGGAAGG
TCGGGATGAATACGACGAAGTGGCCATGCCTGTG**TGA**GGGGCGGGCCCAGGCCCTAGACCCTCCCCCCACCATTCCAC
ATCCCCACCTTCCAAGGAAAAGCAGAAGTTCATGGGCACCTCATGGACTCCAGGATCCTCCTGGAGCAGCAGCTGAG
GCCCCAGGGCTGTGGGTGGGGAAGGAAGGCGTGTCCAGGAGACCTTCCACAAAGGGTAGCCTGGCTTTTCTGGCTGG
GGATGGCCGATGGGGCCCACATTAGGGGGTTTGTTGCACAGTCCCTCCTGTTGCCACACTTTCACTGGGGATCCCGT
GCTGGAAGACTTAGATCTGAGCCCTCCCTCTTCCCAGCACAGGCAGGGGTAGAAGCAAAGGCAGGAGGTGGGTGAGC
GGGTGGGGTGCTTGCTGTGTGACCTTGGGCAAGTCCCTTGACCTTTCCAGCCTATATTTCCTCTTCTGTAAAATGGG
TATATTGATGATAATACCCACATTACAGGATGGTTACTGAGGACCAAAGATACATGTAAAATAGGGCTTTGTAAACT
CCACAGGGACTGTTCTATAGCAGTCATCATTTGTCTTTGAACGTACCCAAGGTCACATAGCTGGGATTTGAACTGAG
CCGTGCAGCT

FIGURE 11

```
CTCTTTTGCTNGGACTTCACTGTCACTCANGAAAAAGCNGTGAAACTAAAACAGAAGAATCTTAGCACTGAGATAAG
GGAGAACCTGTCAGAGCTCCGTCAGGAGGAATTCCAAGTTGACGTTCAATCAGCTGCTGACCCGCTTCTCTGCCTACA
TGGTAGCCTGGGTTGTCTCTACAGGAGTGGCCATAGCCTGCTGTGCAGCCGTTTATTACCTGGCTGAGTACAACTTA
GAGTTCCTGAAGACACACAGTAACCCTGGGGCGGTGCTGTTACTGCCTTTCGTTGTGTCCTGCATTAATCTGGCCGT
GCCATGCATCTACTCCATGTTCAGGCTTGTGGAGAGGTACGAGATGCCACGGCACGAAGTCTACGTTCTCCTGATCC
GAAACATCTTTTTGAAAATATCAATCATTGGCATTCTTTGTTACTATTGGCTCAACACCGTGGCCCTGTCTGGTGAA
GAGTGTTGGGAAACCCTCATTGGCCAGGACATCTACCGGCTCCTTCTGATGGATTTTGTGTTCTCTTTAGTCAATTC
CTTCCTGGGGGAGTTTCTGAGGAGAATCATTGGGATGCAACTGATCACAAGTCTTGGCCTTCAGGAGTTTGACATTG
CCAGGAACGTTCTAGAACTGATCTATGCACAAACTCTGGTGTGGATTGGCATCTTCTTCTGCCCCCTGCTGCCCTTT
ATCCAAATGATTATGCTTTTCATCATGTTCTACTCCAAAAATATCAGCCTGATGATGAATTTCCAGCCTCCGAGCAA
AGCCTGGCGGGCCTCACAGATGATGACTTTCTTCATCTTCTTGCTCTTTTTCCCATCCTTCACCGGGGTCTTGTGCA
CCCTGGCCATCACCATCTGGAGATTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTCTGCCTCTCTTCATTCAC
TCCATCTACAGCTGGATCGACACCCTAAGTACACGGCCTGGCTACCTGTGGGTTGTTTGGATCTATCGGAACCTCAT
TGGAAGTGTGCACTTCTTTTTCATCCTCACCCTCATTGTGCTAATCATCACCTATCTTTACTGGCAGATCACAGAGG
GAAGGAAGATTATGATAAGGCTGCTCCATGAGCAGATCATTAATGAGGGCAAAGATAAAATGTTCCTGATAGAAAAA
TTGATCAAGCTGCAGGATATGGAGAAGAAAGCAAACCCCAGCTCACTTGTTCTGGAAAGGAGAGAGGTGGAGCAACA
AGGCTTTTTGCATTTGGGGGAACATGATGGCAGTCTTGACTTGCGATCTAGAAGATCAGTTCAAGAAGGTAATCCAA
GGGCCTGATGACTCTTTTGGTAACCAGACACCAATCAAATAAGGGGAGGAGATGAAAATGGAATGATTTCTTCCATG
CCACCTGTGCCTTTAGGAACTGCCCAGAAGAAAATCCAAGGCTTTAGCCAGGAGCGGAAACTGACTACCATGTAATT
ATCAAAGTAAAATTGGGCATTCCATGCTATTTTTAATACCTGGATTGCTGATTTTTCAAGACAAAATACTTGGGGTT
TTCCAATAAAGATTGTTGTAATATTGAAATGAGCCTACAAAAACCTAGGAAGAGATAACTAGGGAATAATGTATATT
ATCTTCAAGAAATGTGTGCAGGAATGATTGGTTCTTAGAAATCTCTCCTGCCAGACTTCCCAGACCTGGCAAAGGTT
TAGAAACTGTTGCTAAGAAAAGTGGTCCATCCTGAATAAACATGTAATACTCCAGCAGGGATATGAAGCCTCTGAAT
TGTAGAACCTGCATTTATTTGTGACTTTGAACTAAAGACATCCCCCATGTCCCAAAGGTGGAATACAACCAGAGGTC
TCATCTCTGAACTTTCTTGCGTACTGATTACATGAGTCTTTGGAGTCGGGGATGGAGGAGGTTCTGCCCCTGTGAGG
TGTTATACATGACCATCAAAGTCCTACGTCAAGCTAGCTTTGCACAGTGGCAGTACCGTAGCCAATGAGATTTATCC
GAGACGCGATTATTGCTAATTGGAAATTTTCCCAATACCCCACCGTGATGACTTGAAATATAATCAGCGCTGGCAAT
TTTTGACAGTCTCTACGGAGACTGAATAAGAAAAAAGAAAAGAAAAGAAATTAGCTGGGTGCGATGGCTTATGCCTG
TAATCCCGGCACTTTGGGAGGCTGAGGCAAGCGGATCACTTAATGTCAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCGTCTCTACTAAGGATAAAAAAAACTGGCTGGGCGTGGTGGTACATGCCT
```

FIGURE 12

GCTTCAGGGTACAGCTCCCCCGCAGCCAGAAGCCGGGCCTGCAGCCCCTCAGCACCGCTCCGGGACACCCCACCCGC
TTCCCCAGGCGTGACCTGTCAACAGCAACTTCGCGGTGTGGTGAACTCTCTGAGGAAAAACCATTTTGATTATTACTC
TCAGACGTGCGTGGCAACAAGTGACTGAGACCTAGAAATCCAAGCGTTGGAGGTCCTGAGGCCAGCCTAAGTCGCTT
CAAA**ATG**GAACGAAGGCGTTTGTGGGGTTCCATTCAGAGCCGATACATCAGCATGAGTGTGTGGACAAGCCCACGGA
GACTTGTGGAGCTGGCAGGGCAGAGCCTGCTGAAGGATGAGGCCCTGGCCATTGCCGCCCTGGAGTTGCTGCCCAGG
GAGCTCTTCCCGCCACTCTTCATGGCAGCCTTTGACGGGAGACACAGCCAGACCCTGAAGGCAATGGTGCAGGCCTG
GCCCTTCACCTGCCTCCCTCTGGGAGTGCTGATGAAGGGACAACATCTTCACCTGGAGACCTTCAAAGCTGTGCTTG
ATGGACTTGATGTGCTCCTTGCCCAGGAGGTTCGCCCCAGGAGGTGGAAACTTCAAGTGCTGGATTTACGGAAGAAC
TCTCATCAGGACTTCTGGACTGTATGGTCTGGAAACAGGGCCAGTCTGTACTCATTTCCAGAGCCAGAAGCAGCTCA
GCCCATGACAAAGAAGCGAAAAGTAGATGGTTTGAGCACAGAGGCAGAGCAGCCCTTCATTCCAGTAGAGGTGCTCG
TAGACCTGTTCCTCAAGGAAGGTGCCTGTGATGAATTGTTCTCCTACCTCATTGAGAAAGTGAAGCGAAAGAAAAAT
GTACTACGCCTGTGCTGTAAGAAGCTGAAGATTTTTGCAATGCCCATGCAGGATATCAAGATGATCCTGAAAATGGT
GCAGCTGGACTCTATTGAAGATTTGGAAGTGACTTGTACCTGGAAGCTACCCACCTTGGCGAAATTTTCTCCTTACC
TGGGCCAGATGATTAATCTGCGTAGACTCCTCCTCTCCCACATCCATGCATCTTCCTACATTTCCCCGGAGAAGGAA
GAGCAGTATATCGCCCAGTTCACCTCTCAGTTCCTCAGTCTGCAGTGCCTGCAGGCTCTCTATGTGGACTCTTTATT
TTTCCTTAGAGGCCGCCTGGATCAGTTGCTCAGGCACGTGATGAACCCCTTGGAAACCCTCTCAATAACTAACTGCC
GGCTTTCGGAAGGGGATGTGATGCATCTGTCCCAGAGTCCCAGCGTCAGTCAGCTAAGTGTCCTGAGTCTAAGTGGG
GTCATGCTGACCGATGTAAGTCCCGAGCCCCTCCAAGCTCTGCTGGAGAGAGCCTCTGCCACCCTCCAGGACCTGGT
CTTTGATGAGTGTGGGATCACGGATGATCAGCTCCTTGCCCTCCTGCCTTCCCTGAGCCACTGCTCCCAGCTTACAA
CCTTAAGCTTCTACGGGAATTCCATCTCCATATCTGCCTTGCAGAGTCTCCTGCAGCACCTCATCGGGCTGAGCAAT
CTGACCCACGTGCTGTATCCTGTCCCCCTGGAGAGTTATGAGGACATCCATGGTACCCTCCACCTGGAGAGGCTTGC
CTATCTGCATGCCAGGCTCAGGGAGTTGCTGTGTGAGTTGGGGCGGCCCAGCATGGTCTGGCTTAGTGCCAACCCCT
GTCCTCACTGTGGGGACAGAACCTTCTATGACCCGGAGCCCATCCTGTGCCCCTGTTTCATGCCTAAC**TAG**CTGGGT
GCACATATCAAATGCTTCATTCTGCATACTTGGACACTAAAGCCAGGATGTGCATGCATCTTGAAGCAACAAAGCAG
CCACAGTTTCAGACAAATGTTCAGTGTGAGTGAGGAAAAACATGTTCAGTGAGGAAAAAACATTCAGACAAATGTTCA
GTGAGGAAAAAAAGGGGAAGTTGGGGATAGGCAGATGTTGACTTGAGGAGTTAATGTGATCTTTGGGGAGATACATC
TTATAGAGTTAGAAATAGAATCTGAATTTCTAAAGGGAGATTCTGGCTTGGGAAGTACATGTAGGAGTTAATCCCTG
TGTAGACTGTTGTAAAGAAACTGTTGAAAATAAAGAGAAGCAATGTGAAGCAAAAAAAAAAAAAAAAAAAA

FIGURE 13

```
GGGCAGGCTCAGTGTGAGTGAACTGGAGGCTTCTCTACAACATGACCCAAAGGAGCATTGCAGGTCCTATTTGCAAC
CTGAAGTTTGTGACTCTCCTGGTTGCCTTAAGTTCAGAACTCCCATTCCTGGGAGCTGGAGTACAGCTTCAAGACAA
TGGGTATAATGGATTGCTCATTGCCATTAATCCTCAGGTACCTGAGAATCAGAACCTCATCTCAAACATTAAGGAAA
TGATAACTGAAGCTTCATTTACCTATTTAATGCTACCAAGAGAAGAGTATTTTTCAGAAATATAAAGATTTTAATA
CCTGCCACATGGAAAGCTAATAATAACAGCAAAATAAAACAAGAATCATATGAAAAGGCAAATGTCATAGTGACTGA
CTGGTATGGGGCACATGGAGATGATCCATACACCCTACAATACAGAGGGTGTGGAAAAGAGGGAAAATACATTCATT
TCACACCTAATTTCCTACTGAATGATAACTTAACAGCTGGCTACGGATCACGAGGCCGAGTGTTTGTCCATGAATGG
GCCCACCTCCGTTGGGGTGTGTTCGATGAGTATATCAATGACAAACCTTTCTACATAAATGGGCAAAATCAAATTAA
AGTGACAAGGTGTTCATCTGACATCACAGGCATTTTTGTGTGTGAAAAAGGTCCTTGCCCCCAAGAAAACTGTATTA
TTAGTAAGCTTTTTAAAGAAGGATGCACCTTTATCTACAATAGCACCCAAAATGCAACTGCACTCAATAATGTTCATG
CAAAGTTTATCTTCTGTGGTTGAATTTTGTAATGCAAGTACCCACAACCAAGAAGCACCAAACCTACAGAACCAGAT
GTGCAGCCTCAGAAGTGCATGGGATGTAATCACAGACTCTGCTGACTTTCACCACAGCTTTCCCATGAATGGGACTG
AGCTTCCACCTCCTCCCACATTCTCGCTTGTACAGGCTGGCGACAAAGTGGTCTGTTTAGTGCTGGATGTGTCCAGC
AAGATGGCAGAGGCTGACAGACTCCTTCAACTACAACAAGCCGCAGAATTTTATTTGATGCAGATTGTTGAAATTCA
TACCTTCGTGGGCATTGCCAGTTTCGACAGCAAAGGAGAGATCAGAGCCCAGCTACACCAAATTAACAGCAATGATG
ATCGAAAGTTGCTGGTTTCATATCTGCCCACCACTGTATCAGCTAAAACAGACATCAGCATTTGTTCAGGGCTTAAG
AAAGGATTTGAGGTGGTTGAAAAACTGAATGGAAAAGCTTATGGCTCTGTGATGATATTAGTGACCAGCGGAGATGA
TAAGCTTCTTGGCAATTGCTTACCCACTGTGCTCAGCAGTGGTTCAACAATTCACTCCATTGCCCTGGGTTCATCTG
CAGCCCCAAATCTGGAGGAATTATCACGTCTTACAGGAGGTTTAAAGTTCTTTGTTCCAGATATATCAAACTCCAAT
AGCATGATTGATGCTTTCAGTAGAATTTCCTCTGGAACTGGAGACATTTTCCAGCAACATATTCAGCTTGAAAGTAC
AGGTGAAAATGTCAAACCTCACCATCAATTGAAAAACACAGTGACTGTGGATAATACTGTGGGCAACGACACTATGT
TTCTAGTTACGTGGCAGGCCAGTGGTCCTCCTGAGATTATATTATTTGATCCTGATGGACGAAAATACTACACAAAT
AATTTTATCACCAATCTAACTTTTCGGACAGCTAGTCTTTGGATTCCAGGAACAGCTAAGCCTGGGCACTGGACTTA
CACCCTGAACAATACCCATCATTCTCTGCAAGCCCTGAAAGTGACAGTGACCTCTCGCGCCTCCAACTCAGCTGTGC
CCCCAGCCACTGTGGAAGCCTTTGTGGAAAGAGACAGCCTCCATTTTCCTCATCCTGTGATGATTTATGCCAATGTG
AAACAGGGATTTTATCCCATTCTTAATGCCACTGTCACTGCCACAGTTGAGCCAGAGACTGGAGATCCTGTTACGCT
GAGACTCCTTGATGATGGAGCAGGTGCTGATGTTATAAAAAATGATGGAATTTACTCGAGGTATTTTTTCTCCTTTG
CTGCAAATGGTAGATATAGCTTGAAAGTGCATGTCAATCACTCTCCCAGCATAAGCACCCCAGCCCACTCTATTCCA
GGGAGTCATGCTATGTATGTACCAGGTTACACAGCAAACGGTAATATTCAGATGAATGCTCCAAGGAAATCAGTAGG
CAGAAATGAGGAGGAGCGAAAGTGGGGCTTTAGCCGAGTCAGCTCAGGAGGCTCCTTTTCAGTGCTGGGAGTTCCAG
CTGGCCCCCACCCTGATGTGTTCCACCATGCAAAATTATTGACCTGGAAGCTGTAAAAGTAGAAGAGGAATTGACC
CTATCTTGGACAGCACCTGGAGAAGACTTTGATCAGGGCCAGGCTACAAGCTATGAAATAAGAATGAGTAAAAGTCT
ACAGAATATCCAAGATGACTTTAACAATGCTATTTTAGTAAATACATCAAAAGCGAAATCCTCAGCAAGCTGGCATCA
GGGAGATATTTACGTTCTCACCCCAAATTTCCACGAATGGACCTGAACATCAGCCAAATGGAGAAACACATGAAAGC
CACAGAATTTATGTTGCAATACGAGCAATGGATAGGAACTCCTTACAGTCTGCTGTATCTAACATTGCCCAGGCGCC
TCTGTTTATTCCCCCCAATTCTGATCCTGTACCTGCCAGAGATTATCTTATATTGAAAGGAGTTTTAACAGCAATGG
GTTTGATAGGAATCATTTGCCTTTATTATAGTTGTGACACATCATACTTTAAGCAGGAAAAAAGAGAGCAGACAAGAAA
GAGAATGGAACAAAATTATTATAAATAAATATCCAAAGTGTCTTCCTTCTTTAGATATAAGACCCATGGCCTTCGACT
ACAAAAACATACTAACAAAGTCAAATTAACATCAAAACTGTATTAAAATGCATTGAGTTTTTGTACAATACAGATAA
GATTTTTACATGGTAGATCAACAATTCTTTTTGGGGGTAGATTAGAAAACCCTTACACTTTGGCTATGAACAAATAA
TAAAAATTATTCTTTAAAGTAATGTCTTTAAAGGCAAAGGGAAGGGTAAAGTCGGACCAGTGTCAAGGAAAGTTTGT
TTTATTGAGGTGGAAAAATAGCCCCAAGCAGAGAAAAGGAGGGTAGGTCTGCATTATAACTGTCTGTGTGAAGCAAT
CATTTAGTTACTTTGATTAATTTTTCTTTTCTCCTTATCTGTGCAGTACAGGTTGCTTGTTTACATGAAGATCATGC
TATATTTTATATATGTAGCCCCTAATGCAAAGCTCTTTACCTCTTTGCTATTTTGTTATATATATTTCAGATGACATC
TCCCTGCTAATGCTCAGAGATCTTTTTTCACTGTAAGAGGTAACCTTTAACAATATGGGTATTACCTTTGTCTCTTC
ATACCGGTTTTATGACAAAGGTCTATTGAATTTATTTGTXTGTAAGTTTCTACTCCCATCAAAGCAGCTTTCTAAGT
TTATTGCCTTGGGTTATTATGGAATGATAGTTATAGCCCCXTATAATGCCTTACCTAGGAAA
```

FIGURE 14

GTCATATTGAACATTCCAGATACCTATCATTACTCGATGCTGTTGATAACAGCAAG**ATG**GCTTTGAACTCAGGGTCA
CCACCAGCTATTGGACCTTACTATGAAAACCATGGATACCAACCGGAAAACCCCTATCCCGCACAGCCCACTGTGGT
CCCCACTGTCTACGAGGTGCATCCGGCTCAGTACTACCCGTCCCCCGTGCCCCAGTACGCCCCGAGGGTCCTGACGC
AGGCTTCCAACCCCGTCGTCTGCACGCAGCCCAAATCCCCATCCGGGACAGTGTGCACCTCAAAGACTAAGAAAGCA
CTGTGCATCACCTTGACCCTGGGGACCTTCCTCGTGGGAGCTGCGCTGGCCGCTGGCCTACTCTGGAAGTTCATGGG
CAGCAAGTGCTCCAACTCTGGGATAGAGTGCGACTCCTCAGGTACCTGCATCAACCCCTCTAACTGGTGTGATGGCG
TGTCACACTGCCCCGGCGGGGAGGACGAGAATCGGTGTGTTCGCCTCTACGGACCAAACTTCATCCTTCAGATGTAC
TCATCTCAGAGGAAGTCCTGGCACCCTGTGTGCCAAGACGACTGGAACGAGAACTACGGGCGGGCGGCCTGCAGGGA
CATGGGCTATAAGAATAATTTTTACTCTAGCCAAGGAATAGTGGATGACAGCGGATCCACCAGCTTTATGAAACTGA
ACACAAGTGCCGGCAATGTCGATATCTATAAAAAACTGTACCACAGTGATGCCTGTTCTTCAAAAGCAGTGGTTTCT
TTACGCTGTTTAGCCTGCGGGGTCAACTTGAACTCAAGCCGCCAGAGCAGGATCGTGGGCGGTGAGAGCGCGCTCCC
GGGGGCCTGGCCCTGGCAGGTCAGCCTGCACGTCCAGAACGTCCACGTGTGCGGAGGCTCCATCATCACCCCCGAGT
GGATCGTGACAGCCGCCCACTGCGTGGAAAAACCTCTTAACAATCCATGGCATTGGACGGCATTTGCGGGGATTTTG
AGACAATCTTTCATGTTCTATGGAGCCGGATACCAAGTACAAAAAGTGATTTCTCATCCAAATTATGACTCCAAGAC
CAAGAACAATGACATTGCGCTGATGAAGCTGCAGAAGCCTCTGACTTTCAACGACCTAGTGAAACCAGTGTGTCTGC
CCAACCCCAGGCATGATGCTGCAGCCAGAACAGCTCTGCTGGATTTCCGGGTGGGGGGCCACCGAGGAGAAAGGGAAG
ACCTCAGAAGTGCTGAACGCTGCCAAGGTGCTTCTCATTGAGACACAGAGATGCAACAGCAGATATGTCTATGACAA
CCTGATCACACCAGCCATGATCTGTGCCGGCTTCCTGCAGGGGAACGTCGATTCTTGCCAGGGTGACAGTGGAGGGC
CTCTGGTCACTTCGAACAACAATATCTGGTGGCTGATAGGGGATACAAGCTGGGGTTCTGGCTGTGCCAAAGCTTAC
AGACCAGGAGTGTACGGGAATGTGATGGTATTCACGGACTGGATTTATCGACAAATGAAGGCAAACGGC**TAA**TCCAC
ATGGTCTTCGTCCTTGACGTCGTTTTACAAGAAAACAATGGGGCTGGTTTTGCTTCCCCGTGCATGATTTACTCTTA
GAGATGATTCAGAGGTCACTTCATTTTTATTAAACAGTGAACTTGTCTGGCTTTGGCACTCTCTGCCATACTGTGCA
GGCTGCAGTGGCTCCCCTGCCCAGCCTGCTCTCCCTAACCCCTTGTCCGCAAGGGGTGATGGCCGGCTGGTTGTGGG
CACTGGCGGTCAATTGTGGAAGGAAGAGGGTTGGAGGCTGCCCCCATTGAGATCTTCCTGCTGAGTCCTTTCCAGGG
GCCAATTTTGGATGAGCATGGAGCTGTCACTTCTCAGCTGCTGGATGACTTGAGATGAAAAAGGAGAGACATGGAAA
GGGAGACAGCCAGGTGGCACCTGCAGCGGCTGCCCTCTGGGGCCACTTGGTAGTGTCCCCAGCCTACTTCACAAGGG
GATTTTGCTGATGGGTTCTTAGAGCCTTAGCAGCCCTGGATGGTGGCCAGAAATAAAGGGGACCAGCCCTTCATGGGT
GGTGACGTGGTAGTCACTTGTAAGGGGAACAGAAACATTTTTGTTCTTATGGGGTGAGAATATAGACAGTGCCCTTG
GTGCGAGGGAAGCAATTGAAAAGGAACTTGCCCTGAGCACTCCTGGTGCAGGTCTCCACCTGCACATTGGGTGGGGC
TCCTGGGAGGGAGACTCAGCCTTCCTCCTCATCCTCCCTGACCCTGCTCCTAGCACCCTGGAGAGTGAATGCCCCTT
GGTCCCTGGCAGGGCGCCAAGTTTGGCACCATGTCGGCCTCTTCAGGCCTGATAGTCATTGGAAATTGAGGTCCATG
GGGGAAATCAAGGATGCTCAGTTTAAGGTACACTGTTTCCATGTTATGTTTCTACACATTGATGGTGGTGACCCTGA
GTTCAAAGCCATCTT

FIGURE 15

GGATTTCCGGGCTCC**ATG**GCAAGATCCCTTCTCCTGCCCCTGCAGATCCTACTGCTATCCTTAGCCTTGGAAACTGC
AGGAGAAGAAGCCCAGGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCCACCCATGGCAGGTGGCCC
TGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGGTGCTCACTGCCGCCCACTGCAAG
ATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTC
ATTCCGCCACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGTCGTGAAGCTCAATAGCCAGGCCAGGC
TGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCCTGGAACCACCTGTACTGTCTCCGGCTGG
GGCACTACCACGAGCCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCCAGGA
CTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCCGACTCCAAGAAAAACGCCT
GCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCTTGC
GGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAGCA
TCGC**TAA**CGCCACACTGAGTTAATTAACTGTGTGCTTCCAACAGAAAATGCACAGGAGTGAGGACGCCGATGACCTA
TGAAGTCAAATTTGACTTTACCTTTCCTCAAAGATATATTTAAACCTCATGCCCTGTTGATAAACCAATCAAATTGG
TAAAGACCTAAAACCAAAACAAATAAAGAAACACAAAACCCTCAA

FIGURE 16

```
CCGAGACTCACGGTCAAGCTAAGGCGAAGAGTGGGTGGCTGAAGCCATACTATTTTATAGAATTA**ATG**GAAAGCAGA
AAAGACATCACAAACCAAGAAGAACTTTGGAAAATGAAGCCTAGGAGAAATTTAGAAGAAGACGATTATTTGCATAA
GGACACGGGAGAGACCAGCATGCTAAAAAGACCTGTGCTTTTGCATTTGCACCAAACAGCCCATGCTGATGAATTTG
ACTGCCCTTCAGAACTTCAGCACACACAGGAACTCTTTCCACAGTGGCACTTGCCAATTAAAATAGCTGCTATTATA
GCATCTCTGACTTTTCTTTACACTCTTCTGAGGGAAGTAATTCACCCTTTAGCAACTTCCCATCAACAATATTTTTA
TAAAATTCCAATCCTGGTCATCAACAAAGTCTTGCCAATGGTTTCCATCACTCTCTTGGCATTGGTTTACCTGCCAG
GTGTGATAGCAGCAATTGTCCAACTTCATAATGGAACCAAGTATAAGAAGTTTCCACATTGGTTGGATAAGTGGATG
TTAACAAGAAAGCAGTTTGGGCTTCTCAGTTTCTTTTTTGCTGTACTGCATGCAATTTATAGTCTGTCTTACCCAAT
GAGGCGATCCTACAGATACAAGTTGCTAAACTGGGCATATCAACAGGTCCAACAAAATAAAGAAGATGCCTGGATTG
AGCATGATGTTTGGAGAATGGAGATTTATGTGTCTCTGGGAATTGTGGGATTGGCAATACTGGCTCTGTTGGCTGTG
ACATCTATTCCATCTGTGAGTGACTCTTTGACATGGAGAGAATTTCACTATATTCAGAGCAAGCTAGGAATTGTTTC
CCTTCTACTGGGCACAATACACGCATTGATTTTTGCCTGGAATAAGTGGATAGATATAAAACAATTTGTATGGTATA
CACCTCCAACTTTTATGATAGCTGTTTTCCTTCCAATTGTTGTCCTGATATTTAAAAGCATACTATTCCTGCCATGC
TTGAGGAAGAAGATACTGAAGATTAGACATGGTTGGGAAGACGTCACCCAAAATTAACAAAACTGAGATATGTTCCCA
GTTG**TAG**AATTACTGTTTACACACATTTTTGTTCAATATTGATATATTTTATCACCAACATTTCAAGTTTGTATTTG
TTAATAAAATGATTATTCAAGGAAAAAAAAAAAAAAAAAAA
```

FIGURE 17A

GGTGGAGACACCGCCTCAGGGCTCGGTGCACAGTGGACATTTGGGGAGCGTTGTGGGTGACCCCCACACAGGCACTG
GGAATGCAGGGGAGAGGGGGCCAAGGGGGAAAGGGGCCAGAGTGTTGGCTTTGGATTCAGGAGGG**ATG**GATTCCAGT
CCTAGCTTGCCACTTATTAGGACTCCTGAGAGCAGCCTCCATGAGGCCCTGGACCAGTGCATGACCGCCCTGGACCT
CTTCCTCACCAACCAGTTCTCAGAAGCACTCAGCTACCTCAAGCCCAGAACCAAGGAAAGCATGTACCACTCACTGA
CATATGCCACCATCCTGGAGATGCAGGCCATGATGACCTTTGACCCTCAGGACATCCTGCTTGCCGGCAACATGATGA
AAGGAGGCACAGATGCTGTGTCAGAGGCACCGGAGGAAGTCTTCTGTAACAGATTCCTTCAGCAGCCTGGTGAACCG
CCCCACGCTGGGCCAATTCACTGAAGAGGAAATCCACGCTGAGGTCTGCTATGCAGAGTGCCTGCTGCAGCGAGCAG
CCCTGACCTTCCTGCAGGGTTCCTCACACGGAGGGCAGTCAGGCCCAGAGCCTTGCATGATCCCTCTCACGCCTGC
AGCTGCCCACCTGGGCCAGGCCGTCAGCATCTTTTCCTCCTGCAGGACGAGAACATGGTGAGCTTCATCAAAGGCGG
CATCAAAGTTCGAAACAGCTACCAGACCTACAAGGAGCTGGACAGCCTTGTTCAGTCCTCACAATACTGCAAGGGTG
AGAACCACCCGCACTTTGAAGGAGGAGTGAAGCTTGGTGTAGGGGCCTTCAACCTGACACTGTCCATGCTTCCTACT
AGGATCCTGAGGCTGTTGGAGTTTGTGGGGTTTTCAGGAAACAAGGACTATGGGCTGCTGCAGCTGGAGGAGGGAGC
GTCAGGGCACAGCTTCCGCTCTGTGCTCTGTGTCATGCTCCTGCTGTGCTACCACACCTTCCTCACCTTCGTGCTCG
GTACTGGGAACGTCAACATCGAGGAGGCCGAGAAGCTCTTGAAGCCCTACCTGAACCGGTACCCTAAGGGTGCCATC
TTCCTGTTCTTTGCAGGGAGGATTGAAGTCATTAAAGGCAACATTGATGCAGTGAGTGATGGGGGTCCGGGCCGGGG
CTGGGGATCCCTCGGGGTCTCCCAGACCAGCAGGAAGTCAGGCACATGTGACATACTCAGGGACAGGATAGACTGGG
GGCGGGGGGGGGGCCAAGAGAGAACCAACCAGAGAGCAGGGGCAGGAGGAGGCCCTTCTGGCAGAGCAGCCTGGGAAG
ACAAGGGAGGAGGAGGCATTTGTGGTGCCTGGGATTTTGACTGGGAGATATAGGACTGCAGCATTGCAGTGGAGGGA
GGTGAGGGAGGTGCT**TGA**GGGAGGCAGAGGTTAGGAAAGCCCATCTGTTTAGGGCATGACGATTAGGCTGGAGTCT
GGTACCTCCCCTCCATTATAGCTCTCTCCTGCTCTTTCATTTTGTTACTAAAAACCAGAGTCCTAGGCGGGGGCTGT
ATTTGAGCCCAACGTCATGTAAGACTTAGGAGGTAAAACCAGGACTGGAGGCCAGATCTCCTGGCTCCTGGGGCCCC
ACCTGAGCCTAGCACAGGGCTGGACCACTATGCCCTGGAGGAGTCCCGGTCTGCTGTGGTGTTGGGAGGTTCGGAGG
ATGCAGAGGGGTTGGGGCTGGGTGGGCACCCGTCAGGCTGACCAGAAGGTGCCTGCAGGCCATCCGGCGTTTCGAGG
AGTGCTGTGAGGCCCAGCAGCACTGGAAGCAGTTCCACCACATGTGCTACTGGGAGCTGATGTGGTGCTTCACCTAC
AAGGGCCAGTGGAAGATGTCCTACTTCTACGCCGACCTGCTCAGCAAGGAGAACTGCTGGTCCAAGGTGGGCTGATG
CCACGTGTTAGGGGCATTGGGTGACCAGGGCTGACTGTGTGCCTCCAGACCACGGGCCAAATCCCTAACTGAACACA
GATGTCTCAGCTGGAATCTAAACATAACCTTAAATTCTAATAGGACTCAGGCTTGGAAGGAACTAAAGACCACAAGA
GAAACTTCTGAACCACAATGTGTCACAAAGAGATTTCTAGTCACAAAGGACAGAAACATGGCTCCCTCTGTCCAGTA
GAACTGTTTCTGTGGTAGAAATGTTCTACTTATGCACTGCCCAAATATGGTAGTCACCAGCCACATGTGCACAATGA
GCACATGAAATGTGCCTAGTGCAACTGGGAAACTGATTGTTTTCAGTTTTATTTAATTTTAATTAACTAAATGTTAA
ATTTAAATAGCCATGTAGGGCTTGTGGCCACTATATTGGACTATGCAGGTCCAAAACACAAAAGGCTCATATAACTG
AACATTCTTGGCACATCCGACTTCAGGTAGGGCTGGATCCAGGAATTCAAATGATGTCGTCTGGCTTGGTCTTTCCA
TTTGTGGCTGTGCTCTCTCCTATGACATCTTTGTTTCTGCTGCATCTTTGCAGGAAGGTTCTCTCCATGTGACAGGC
AAGGTGGCCACGGGCTGCTTCTACTCATATCCTCCCCTTGGTTTCAACCAGAGTCCCACAGTTTGAGCCTCACTGGT
CTGACTTGTCACCTGCCAATGCCTGGAACAGAGGGGTGGGAAGATTCCTCATGGGCTGAGAGTAGGAGGGGGTGGT
TCCACAGAAGAAAATGATGCACCCAGAAATAGTGGAGGGATTAACAAGATGCCATACAGGCAAAACAAAGCCAAACA
GATGCCCGCCTACCAAAGATGAAATTTATCATGGTAAGTATTGAAATAAGTGTTAGCTTGTACCATCATAGTAATGA
TAGTGCAGAAATTGGAACCAAGAGTCTTACAACCACCTAGCTCAGCAAACGTCTAATCTGTTATTTGTAAATACACA
GGACATGTGTCTTCATGGCTTCATCCCCTGCCCATAGCAGACATTGCTAATCAATCCTCTGCCATGAGCCTAGCTGT
GACCTTAGATACCTTCCCTGCAAAGCTCCAGGCAGCCGTTAGAACTGACCCACATTGGCACCTGTTAGGAGACCTGT
CATCTGCATCTGGTCCCTCCCCCCTCCTCTGCATGGAGGTCTCTTCCGTAGCCCTGCCTGGGGACCAAGCTGAAGGGGG
AATTCTCTGGGTCTGGGGCAGGTGGGGCTGGGGAAGGGAGCAGGCTTCTACTGAGCTCTCAACATGTCTGGCACCTG
TCATGTCCTCCACACCACACCCAGTGAGGAGAAGTTATTATCTCCATTGGGCACATGAGGGCCACTGATGCTAAGCG
AGGACTGTGACTTTCCCACCCACATGGCTGTTGGTGACAGACACCGGATTTGAGGACAGTCCTGTGTGACTCAGAGG
CCTCTGCTTCTCTCTTGCTCCTAGAGTCCTCCTGAGGAGTCGGGGCTTGCCCTGAGCCCACCCTGCTGTTGAAGGTG
CTTCCTCAGGCCCAGCTCCCATGGCCCCCCACACCCCCTCCTCATCACCTCCTACTCCCAAAAAGGACAAAAGCCTCAG
GGAACCTTTTTTCTTTTTTTAGAGACAGGGTCTTGCTATGTTGGTCAGGCTAGTCTTGAACTCCTGGGCCCAAGCAA
TCTTCCCACCTCTGCCTCCTAAAGTGCTGGGATTCTGGGCCTTAGGGAACCTTTTTGAACTGAAAGTGACTCTCGAA
GCCTTCCTGTAGAGTGAGGTGGGTGGCTGGGGCTAACCATATGGGGAAGGAGAGACCTGGTGGGGGCACACAGCTGC
TATATAGAGGAACAGAAGGTGGGCCAGGCCTCCCAGCTGTCAGGAAATCTGGCTCAGTCCCCAGACCGCTGTGTGGC
TTTGGGGTGGCCAGTCCCTCTCTCTCTGCAACACTCTCCTCATGAGGACTTCTGTGAAAAATGGGGGTTGTAACCCT
CTGAGTTCTGGGGCCCCACCCATTTCTAATCTCCAGAACTTGGCCACCACAGACTCCACCAACTTCTAG

FIGURE 17B

```
TCCTGGGGCCTGGGCCTCTGGCCATTGCCATAGGCACCACCTGCTCTGTGCAGGCAGCGCCCCCCTCTGCCAGGATC
CTCCGAGGTCAGCTGCTGGGTCTGACCCGCAGACCCTGGCTGAGCGACGGATGAACGGAGTATGCAGACACAGGCTT
GCCTGTCAGCAGATGGGGGACCCCTGCCAGAGTCAGCAGCGGCCCCCATAAGCCTGCCACGCTTGCATTTATTTAGT
ACAGATGTAATGACAAAGGCCTAAAGCAAACTCCATTTGTGGGTAATTAACATTGTCGCCCCCCCAGAAAGAGCAGT
CCTCCGCATGATGATTAAAGGCCAGGTTCCGAGGCCTAAGTAAACCAACTTATCTAGATCAATTCCCTTACTTCTTG
TTATCTACTCTGAGAGAATTCAGCTGCCTTCAGCCAAATCCTTTCCCGAAGCTTTTGCAAAACCTCCGAGCCTTCCA
AGGTTTGCTTCTTTCTGTAATTTTTCTCACCACCCTGACCTATCTCCTGCAGTCAGCCCTGTGGAGGCCTTTGTGTT
TCCCCCAGTGCTGGCAGCCTAGAGGCTGAGATGGCCAGAAACAAGGTGGTGACAGTGGCGTGCTCAGGGCTTGGGAA
ACCCAAGGAGCTAAAGGCATGCCCAGGCAACCAAAGAGGACAGGAAGGCTTCTGAGGAGAGACCTCTGAGGTGGGTC
TTGGAGAGGAAGGACTTAGGGAGGCAGAGTGGAGGAAGTGAGAGGACACCCCAAGCCAAGAGGGCGGCAGGACCAAA
GGCTCAGAAGCCAGGGCGCTGCAGAGGGGCTGTGTGCCACAGGGTGAAGAGTTTGTGTGGCAGAAGGGCAGGGGGCT
TGCATCAGGGGTGACAGCTGCTCTTTTGTCCCAGCATAGCCCCTGTACATCCCTGGAGAGCTGGGGCGTCCACAACT
CTAAGTCACAGCCCCCATCCTAACCCTGGTGGTGCAGTGAGGGTGAGCTGTCTGTGGGCAGGAGGGAAGACTCTTGG
AGATGAGCCTGGTGAAGGGATAATGGCATCCCGGGCCGAGGAGCAGCACAGGCAGAGGCCTGGGGAGAGTTTAAGGA
GTGTAGGGGAGGAAATGGCAGAAGATGAGCCAGAAAAAGAAAGGTTAGGGCAGGTCCTGGAGGACATGAGTGGCTGT
TTGGGCTTTATCCAGCAGTGGGGGAGCCTTGGCAGGCTTGTGGCTTAGATAGGTGCTTTAGAAAGCCCACCAGCAGT
TGCTGGGCCACCCCGCTGGCTGGGTCCTGTTCTAAGGCAGGAAATACAAGCATGAGCAGGAAAAGACCCCCTCAAGG
CTCACGTCCTAGTGGGGAGACAAGAAACACAGATGGGCAATATAACACGATGTCTGGTTCCAGTAAGTGCAGTGAAG
AACAAGCGAGGCTGGATGCAGGGAGTGATGGGAGGGGCTTTGTAAGGGGAGGTCGGGGGAAGCCTGTCTCAGAGGAC
ACCAGAATGGAGCGCAGGAGCAGCACGTGGCAGTCACATGGCAGGCCGTTAGGGCAGAGGGAGCTGGGCAGGGCACA
GCAGGGCAGGAGTGTGTTTGATGTGTCCTGGGAACCGCCCTGAGGCCGTCGTGTGGCTGGAGTGCTGCAGGTGTCAA
GGAAATTGTAGGAGATGTCTCCTGAGCTGTGATGGAATATAACCAGATTTCCAGAAGGAACTGACATGATCTGACTTA
AAAAGGTCAGTGTGCGAAATGGCTTGCAGGGGACAGGAGTGGGAGCAGGGAGATAGGAGACAATGTGTACCAGGACA
GCAGAAAGACATCCCGGGTAGCCTGGAACAGGGAGACGGTGTGGAGATGGTGGCAGTCCGATAATGAGAGCCGTAGG
GCAAGGCCAGCAGGATCCTAGAGTGAGACGGGAGGTAAAGTCACCGGGACTTGGTGTCTCCACGTCAGGGGCAGGGG
AAAGGGAGAGGACAAGGGTGACCCGGGAGGTTAAAGATGGGACCGGGGCCCAGACGCAGTGGCTCATGCCTGTAATCC
TAGCACTTTGGGAGGCTGAGGCGGGCGGATAGCTTGAGGTCAGGAGTTTGAAACCAGCCTGGCCAACATGGTGAAAC
CCCGTCTCTACTAAAAATATACAAAAATTAGCCTGGCGTGGTGGTGCATGCCTGCAGTCCCAGCTATTCAGGAGGCTG
AGGCAACAAGAATCGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCATAGCACTCCAGCCTTAG
CCTGGGCGACAGAGCGAGACCACATC
```

FIGURE 18

GAGCAGCGCGGCCTGACGGGACCAAGGCGGCGGGAGTCTGCGGTCGGTTCCCTCGGCTGTGGACCGGGCGGCACGCAC
GCGGTGCAGGGTAAC**ATG**GCGGATGCGGAAGTAATTATTTTGCCAAAGAAACATAAGAAGAAAAAGGAGCGGAAGTC
ATTGCCAGAAGAAGATGTAGCCGAAATACAACACGCTGAAGAATTTCTTATCAAACCTGAATCCAAAGTTGCTAAGT
TGGACACGTCTCAGTGGCCCCTTTTGCTAAAGAATTTTGATAAGCTGAATGTAAGGACAACACACTATACACCTCTT
GCATGTGGTTCAAATCCTCTGAAGAGAGAGATTGGGGACTATATCAGGACAGGTTTCATTAATCTTGACAAGCCCTC
TAACCCCTCTTCCCATGAGGTGGTAGCCTGGATTCGACGGATACTTCGGGTGGAGAAGAACAGGGCACAGTGGTACTC
TGGATCCCAAGGTGACTGGTTGTTTAATCGTGTGCATAGAACGAGCCACTCGCTTGGTGAAGTCACAACAGAGTGCA
GGCAAAGAGTATGTGGGGATTGTCCGGCTGCACAATGCTATTGAAGGGGGGACCCAGCTTTCTAGGGCCCTAGAAAC
TCTGACAGGTGCCTTATTCCAGCGACCCCCCACTTATTGCTGCAGTAAAGAGGCAGCTCCGAGTGAGGACCATCTACG
AGAGCAAAATGATTGAATACGATCCTGAAAGAAGATTAGGAATCTTTTGGGTGAGTTGTGAGGCTGGCACCTACATT
CGGACATTATGTGTGCACCTTGGTTTGTTATTGGGAGTTGGTGGTCAGATGCAGGAGCTTCGGAGGGTTCGTTCTGG
AGTCATGAGTGAAAAGGACCACATGGTGACAATGCATGATGTGCTTGATGCTCAGTGGCTGTATGATAACCACAAGG
ATGAGAGTTACCTGCGGCGAGTTGTTTACCCTTTGGAAAAGCTGTTGACATCTCATAAACGGCTGGTTATGAAAGAC
AGTGCAGTAAATGCCATCTGCTATGGGGCCAAGATTATGCTTCCAGGTGTTCTTCGATATGAGGACGGCATTGAGGT
CAATCAGGAGATTGTGGTTATCACCACCAAAGGAAGCAATCTGCATGGCTATTGCATTAATGACCACAGCGGTCA
TCTCTACCTGCGACCATGGTATAGTAGCCAAGATCAAGAGAGTGATCATGGAGAGAGACACTTACCCTCGGAAGTGG
GGTTTAGGTCCAAAGGCAAGTCAGAAGAAGCTGATGATCAAGCAGGGCCTTCTGGACAAGCATGGGAAGCCCACAGA
CAGCACACCTGCCACCTGGAAGCAGGAGTATGTTGACTACAGTGAGTCTGCCAAAAAAGAGGTGGTTGCTGAAGTGG
TAAAAGCCCCGCAGGTAGTTGCCGAAGCAGCAAAACTGCGAAGCGGAAGCGAGAGAGTGAGAGTGAAAGTGACGAG
ACTCCTCCAGCAGCTCCTCAGTTGATCAAGAAGGAAAAGAAGAAGAGTAAGAAGGACAAGAAGGCCAAAGCTGGTCT
GGAGAGCGGGGCCGAGCCTGGAGATGGGGACAGTGATACCACCAAGAAGAAGAAGAAGAAGAAAGCAAAAGAGG
TAGAATTGGTTTCTGAG**TAGT**GAAGGCCACTTGAAGCTGGAGGAGAAACTAAAGCCTTATTGAGAAAACATGTTATA
GATCCTTTTGTTGCTGAGAGAGTGGAACATAGGTCCTAGCACAGGGTGAAGAGTTCTGGCACATTTTAGCTGCTACTT
TGAGACCTCGGTGATGTTACCTGGTGTGGTCATCCCATCTTGTCCTGTTTTAAGGATATGGGTGGTGAAAGATGAAA
GAGGCAGAGTTTATCCCAATGACTTCTCTGTTTGAGTTGGGAAGCCTCACCTTCAGACCCAGTAACTGTCCGCAGCT
GTCTGCTAGTGGTTGTCTTAACATCGTAGTCCTAGTTTGCATTTTTTAAATCCCCTCTGTTTAAAAGGTTTGTAAAA
CAAAAACAAAAAACTAAGTCTGCTCAGTGAAATGCTGTAGAACCCTAAATAAGTGGTAGAAGAGTGTCACTGAATTT
TGTCTCTGAATTCAGTATAACTGAGTTTTGTCCATGCTGGTGTCTGGGTTATAGGCCTGATGGGCCTGGTAGTTTTC
CATCTTGTTCTGGCCTAGAGGTCAGTCCTTTGCACTTCCTCAAAGCTTGTGTACAGTGCTCACCTAAATCCATCTGA
CTACTTGTTCCTGTGCCCTCTTGTTTTAGGCCTCGTTTACTTTTAAAAAATGAAATTGTTCATTGCTGGGAGAAGAA
TGTTGTAATTTTTACTTATTAAAGTCAACTTGTTAAGTTTTTTATGTATTCCTGTTGGGTTTTCTTGTTGATCTCAT
GCTAGCAGAGCAAAAATTGTAAAATATTTTGATTAAAAATCTAGGGACCTTTATGTCCTATTTGGAATTCGATATCAA

.

FIGURE 19

```
GGCACGAGGCTACAAAGCAGGAAAGTATGCTTGGGAGAGGCCAAGTGAGTGGGGAATCAGCCCAAAGCCAGGCGTCC
AGGGTCTCCCTCACCTGAAGCTGACTTTTTCCCCACCTTGGACAGAGGGCGGGAGATGCCATCCCCACTGAACCCAG
TGCTTTCACCAGCCATATTAGCTCCCACTCACCCCCCGTCGTGGAAGCCTCGGCCGTCACACCTGCAGGGCCGGGGC
GTGCATGGCCTCAGGGATGGCCTGTTCAGCTGCTGGGTGACTCGGGTCCAGGTGCCTCACCACCTGCTGAGCTCTGT
GTGATTTCTGGACGCTTCTGCTCGTTGCCTTTGGGCTCAGTGAAGAGTCTGGAGTTTATCTGGAGTGAGGTGGCCGG
TTCTTGGTGGGATCTGAGCAGGACAGCGTCTGGCTCCTTCCCTCGGCTCATGGCCCTCAGAATCTGCGTCACATACA
CCCCAGCTCTCCCGATAGGTCTCTGCACTCGCTGTTGCCTCTGCCTGGAACAGTCTCCCTCCTGGTGTCATTGTCTC
CGTGGTGTGTCCTTCCTGACCTTCCACCTCCACCAGTCTGTCCCCCTTGGGGACAGGGACTCGTTGCTCATGTTCAC
CCGGCAGGCTGGACACTTCGTGGAGGGCTCCAAAGCTGGCAGATCCCGGGGCCGCCTCTGTCTCTCCCAGGCCCTGC
GTGTTGCGGTGAGAGGAGCATTTGTGTCTCTGTGGTTTGCTGCTGGAGCTGGTGACCGGGAGAGAAACAAGGGAGAC
AAGGGTGCCCAGACAGGTGCGGGGCTCAGCCAGGAGGCAGAAGACGTGGACGTGTCCCGGGCCAGGAGGGTCACAGA
TGCACCACAAGGCACTCTGTGTGGCACTGGGAACAGGAATTCTGGGAGTCAGTCTGCAAGGGCGGTGGGCGTTGCTC
ACCTGGGAGAAGCCTTTAGAGTGGGCGTTGAGCAGGCCATTAGCTCGTGCCCTGAGGAGGTGCATGGGCGGCATGGG
CTCTCCATGGAAATTATGTGGGCGCGAATGGATGTGGCTCTGCGCTCACCTGGGCGAGGACTTCTGGCCGGTGCCGG
GGCACTCTGCGTGACCCTGGCAGAATCGAGCTGCCCTGACTATGAAAGGGGAAGAAGAGCATGCCTGACCCTCCACC
GGCACCCCACCCCTCACTGCTCCACCTGGGGCCTGCCTCTGCGGGTGGCTGGGTCCTGGCTGACTGTTGTGACTGTT
GAGGCCCTGGGGGGGTGGCGCATGGGAGTTAGGAGGACTGGCCAGGTGGGGCCCACTATGCACCCACCCCCCAGTGTC
AGGTGCTTCTCCTCTCCTCCTCCACCACCTCCTCCTCCTCCTCCTCATCATCATCCTCACTTGTTGAGGACGTCCTG
TGTGCCAAGTGGTTTATATGCCCAGCCTCATTTAATCCTCAGAATGACTCCATGAGGTAGCTACTAAAACCCCCCAC
TTAACAGATGAGGAAACTGAGGCCTAGAGAAGCTCAATAAGTTGCCTAAGTTCCAAGTTTCCTCTCCCAACTCTCCT
ACCCTCTCCTCTTCCTTCTCCTTTCTCCCATTCTTCCCTGCCTCTTCCCTAACTAGACAATTTTTTATTGAGTGTCT
CCCAGGTGCAGGCATGAGCCAGGTGCTGGGAAAATCATGATAACCCAGCTCCTTCTGGTCATTTTCTCAGCTGGTTA
GAGGCTGGGAGGACACGCAAGTTCAGCTCCAGCCGACTGGGGCATTGGTGGTAGCCCCTGGAGACATTGTGCAATGG
GGCTACGAGGCTGCATCTGGCTCCAGGGAAGCGTGTTGCAATCCATGAGTGATGTCTGCCATGCGTACAGGCATGGA
GAGTGAGGCGCCTGTACTGTCTTTCTGTAGACCCTAGACTGGTGGGGCCTCTGAAATGCATCCAGACACTGTGCTGG
GTGTGTTGCATGGCCCTCCCAACCAATTCAGTATTTTTCTCCCCATTTTCCAGGGAGAAATCTAAGGCGTCAGAATG
TAAGGTTCTTATTGGAACCCAGGCTCCAGGGTCCCTGGTTTTCTGTGACATCATGCTGCAGGACCCTGTTTCTCTCT
TTGCTTTGGCTGCTGGGGAGCTCAGAAGGGAGCTTTAGGCTTGCTCTCAGTCACCACATTAGCTCAGGGGTTTGGGC
ATTTTTGTCGCCGTCTCCTTTGGGCTCTGTCTCCTCCCTGCTGTGCTTCCTGAGGAGCAGGCCGGATGTAAGTTATC
AACTATAAATAAAACCAAGACTTCCGTTCTGGCTCTCAAAAAAAAAAAAAAAAAAAA
```

FIGURE 20

```
GTACAGAAGCAAAATCAAACCTGCTATTTCAGCACTCCTGTTTTTAACTTGGTGTCTTAGTGCTTGGATTGGTGGG
ATGTTTCGGAATGGGCATTGTCGCCAATTTTCAGGAGTTAGCTGTGCCAGTGGTTCATGACGGGGGCGCTCTTTTGG
CCTTTGTCTGTGGTGTCGTGTACACGCTCCTACAGTCCATCATCTCTTACAAATCATGTCCCCAGTGGAACAGTCTC
TCGACATGCCACATACGGATGGTCATCTCTGCCGTTTCTTGCGCAGCTGTCATCCCCATGATTGTCTGTGCTTCACT
AATTTCCATAACCAAGCTGGAGTGGAATCCAAGAGAAAAGGATTATGTATATCACGTAGTGAGTGCGATCTGTGAAT
GGACAGTGGCCTTTGGTTTTATTTTCTACTTCCTAACTTTCATCCAAGATTTCCAGAGTGTCACCCTAAGGATATCC
ACAGAAATCAATGGTGATATTTGAAGAAAGAAGAATTCAGTCTCACTCAGTGAATGTCGCAGGCCATTTCTAAAAGT
GCTACAGAGGACAGACAGGGTTTTGAGGCCACCCTGATTATTGGGATGCATCTGCAGCACATCCAGGACTTGAATTT
CATTACGAGTTCCTAATAGTTGTATTTCTAAAGATGTGTTTCCTAGAGAATGTACAGCCTTATGACACTGTAGTGAT
GTTTTATAATTTTCTAAGTAGATTTTTTTATATTAACAAATTCATATACAGAAAAAATAAGGTGTTACAAAAAATG
GAGAGCTCTTATTTTTGTACAGATTCTGTCGTTTTTTTTTTATTTGTGTGAGATTTATGGAAATACACTAAATGAGTA
ATTCAGGTTCAGTACATTTATTACAAAGTGAAATCAGGGGATATTCATTTGTAAATTTTATTCTTAGTGAATGAACT
GTATAATTTTTTTTATCAGGAGAGCACTTATAAAATTCAATTTATAAAGATCATATACCCAAATCATAAAGATTTAG
TTGATACATTAACACTAAGATACTCTGATTTTTAGCGAACTAAACAAAGTGCTTCTACTGAGAGGCCTTTATACCAC
CATGTACAGTAACTCTAAGTGAATACGGAAGACCTTGGTTTTGAAATTCTGCCACCTTGTTTCTCCCTGCTCATGAG
GTCGCACCTTTTGCTCTTGCTGCTAATTGCCCATTCGTAGTGGGTGTAATGCCAGGTGGAATGGTTTCAACAAGTCA
GGTGAAAACCATCCTTTATTGTTGCTGGCACAACTTGATATATAGTCTGACTCAGAACTGAAGCTCACATCTCAAAT
TCATTTCATGCCAGTAAATGTGGCAAAGAGAAGAAAGGCCCAAGAGCGAGACAAGAAGAATGGAGAAGGGGGCAGCC
AAGAAGAACTTCTGGGTTCAGGGTACTGTTTATTTGCTCCTTCTCTTCATGCCTGTGGCTGGATGTCCCACCAACACT
ATAAGAAATATAAGTCAAGCCCTTTGTGTTAAGCAAGAACTACAGACTCCATCTTTTCACCCAAATCATGAATGACC
AATAAAAAGCAAGTTATTCCAGAGGAAGAAGCAGCCCTTGAAATGTTAAGGCTTAAGGCTTGAAAGGTGAAGAGCAGG
AATTCTCTCTTTCAAATCCTAGAGCATAAACCCATGTGTGGCCAAGTGAGATCAGCCCTCAAGGGCACATGCCAAGG
GCAGAGCAGCCCATGTAGACAGCTTCGGAGGGCATGGGGGTGTAGGGAGTTCGGGGTAGCTCCTCATTAACTATTTG
TTGGGTGAGTAAAGGGGTGAGGCTCAGTGGCAGGTACCTCTGCAATGACAAGCTGCCTCCCCTCTATGTGTTTAGCA
TATGTTATTAGAACATGTCCGACACCCCTACCGCTGCCATTTGGGCCCTTTAATAAAGCCAAGTAGAGAAATCTGGC
AATAAAAGGCAAATGTAAGCATGCTTTCTTTAAGACGCATCATAAATGGTTTTCTTTAAGTGAATGGAAGAGTTTGA
CAGAGATACACCTTTGTAAGAAAACATTAAGAATGCTGGCTGGCTGTGGTGGCTCACACCTGTATTCCCAGCACTTT
GGGAGGCCTAGGCAGGAGGATTGCTTGAGCCTGGGACTTCGAGACCAGACTGGGAAACATGGCAAAATCCCATCTCT
ACAACAAAAATACAAAAATTAGCCAAGTGCGGTGGTGTGCCTGTAGTCCTAGTTACTTGGGAGGCTGAGGTGGGAGA
ATCACCTGAGCCCAGGAGGTGGAGGCTGCAGTGAGCCATGCCAATGCACTCCAGTCTGGGCAACAGAGTGAGACCCT
GTCTCAAAAATAAATAAATAAATAAATGAATAAAAGAGAATGCTAATCATTTCTGGGTTCACTGCGACTCACTGTAGT
GCTGGGGATCCCCCTTGTAACACTGGAACTGAAAGGCAGTGATGAAAGCTATGTCAAGCATTCATTATTCTGAAGAG
GAGGAGAAATGCCACATACCTTTCCCATGGGACCTGTGGTGGAATGAATCCATACTTCTGCCTCACTTCGAGCAGAC
TTTTGTTCTCGGCGCTCCTCACGATGGAGTTTCATGCTTCATTTTCACATCTCTCTGCACAATTAGATTGGGAGCTC
CTTGAGGGCAGAGTACGTGCCTTAATCTTTATCTTTGTAATGCCCACAATGAACAGAGTGCCTCCTGGTACACTGTAG
GAGCTTAAGAAATACTCACTGAATGCATGAATGAATGAATGAACAAATGAAGGAATGACTAAGGATGTTTGTAGTGC
TATAATATAGAATGGGATTTACTCTGCTTTACCAGTTAGTTTCATAATAAACAAATAGTCTGT
```

FIGURE 21

GGCACCGATTCGGGGCCTGCCCGGACTTCGCCGCACGCTGCAGAACCTCGCCCAGCGCCCACC**ATG**CCCCGGCAGCT
CAGCGCGGCGGCCGCGCTCTTCGCGTCCCTGGCCGTAATTTTGCACGATGGCAGTCAAATGAGAGCAAAAGCATTTC
CAGAAACCAGAGATTATTCTCAACCTACTGCAGCAGCAACAGTACAGGACATAAAAAAACCTGTCCAGCAACCAGCT
AAGCAAGCACCTCACCAAACTTTAGCAGCAAGATTCATGGATGGTCATATCACCTTTCAAACAGCGGCCACAGTAAA
AATTCCAACAACTACCCCAGCAACTACAAAAAACACTGCAACCACCAGCCCAATTACCTACACCCTGGTCACAACCC
AGGCCACACCCAACAACTCACACACAGCTCCTCCAGTTACTGAAGTTACAGTCGGCCCTAGCTTAGCCCCTTATTCA
CTGCCACCCACCATCACCCCACCAGCTCATACAGCTGGAACCAGTTCATCAACCGTCAGCCACACAACTGGGAACAC
CACTCAACCCAGTAACCAGACCACCCTTCCAGCAACTTTATCGATAGCACTGCACAAAAGCACAACCGGTCAGAAGC
CTGATCAACCCACCCATGCCCCAGGAACAACGGCAGCTGCCCACAATACCACCCGCACAGCTGCACCTGCCTCCACG
GTTCCTGGGCCCACCCTTGCACCTCAGCCATCGTCAGTCAAGACTGGAATTTATCAGGTTCTAAACGGAAGCAGACT
CTGTATAAAAGCAGAGATGGGGATACAGCTGATTGTTCAAGACAAGGAGTCGGTTTTTTCACCTCGGAGATACTTCA
ACATCGACCCCAACGCAACGCAAGCCTCTGGGAACTGTGGCACCCGAAAATCCAACCTTCTGTTGAATTTTCAGGGC
GGATTTGTGAATCTCACATTTACCAAGGATGAAGAATCATATTATATCAGTGAAGTGGGAGCCTATTTGACCGTCTC
AGATCCAGAGACAGTTTACCAAGGAATCAAACATGCGGTGGTGATGTTCCAGACAGCAGTCGGGCATTCCTTCAAGT
GCGTGAGTGAACAGAGCCTCCAGTTGTCAGCCCACCTGCAGGTGAAAACAACCGATGTCCAACTTCAAGCCTTTGAT
TTTGAAGATGACCACTTTGGAAATGTGGATGAGTGCTCGTCTGACTACACAATTGTGCTTCCTGTGATTGGGGCCAT
CGTGGTTGGTCTCTGCCTTATGGGTATGGGTGTCTATAAAATCCGCCTAAGGTGTCAATCATCTGGATACCAGAGAA
TC**TAA**TTGTTGCCCGGGGGGGAATGAAAATAATGGAATTTAGAGAACTCTTTCATCCCTTCCAGGATGGATGTTGGGA
AATTCCCTCAGAGTGTGGGTCCTTCAAACAATGTAAACCACCATCTTCTATTCAAATGAAGTGAGTCATGTGTGATT
TAAGTTCAGGCAGCACATCAATTTCTAAATACTTTTTGTTTATTTTATGAAAGATATAGTGAGCTGTTTATTTTCTA
GTTTCCTTTAGAATATTTTAGCCACTCAAAGTCAAATTTGAGATATGTTGAATTAACATAATATATGTAAAGTAGA
ATAAGCCTTCAAATTATAAACCAAGGGTCAATTGTAACTAATACTACTGTGTGTGCATTGAAGATTTTATTTTACCC
TTGATCTTAACAAAGCCTTTGCTTTGTTATCAAATGGACTTTCAGTGCTTTTACTATCTGTGTTTTATGGTTTCATG
TAACATACATATTCCTGGTGTAGCACTTAACTCCTTTTCCACTTTAAATTTGTTTTTGTTTTTTGAGACGGAGTTTC
ACTCTTGTCACCCAGGCTGGAGTACAGTGGCACGATCTCGGCTTATGGCAACCTCCGCCTCCCGGGTTCAAGTGATT
CTCCTGCCTTCAGCTTCCCGAGTAGCTGGGATTACAGGCACACACTACCACGCCTGGCTAATTTTTGTATTTTTATTA
TAGACGGGGTTTCACCATGTTGGCCAGACTGGTCTTGAACTCTTGACCTCAGGTGATCCACCCACCTCAGCCTCCCAA
AGTGCTGGGATTACAGGCATGAGCCATTGCGCCCGGCCTTAAATGTTTTTTTTAATCATCAAAAAGAACAACATATC
TCAGGTTGTCTAAGTGTTTTTATGTAAAACCAACAAAAAGAACAAATCAGCTTATATTTTTTATCTTGATGACTCCT
GCTCCAGAATTGCTAGACTAAGAATTAGGTGGCTACAGATGGTAGAACTAAACAATAAGCAAGAGACAATAATAATG
GCCCTTAATTATTAACAAAGTGCCAGAGTCTAGGCTAAGCACTTTATCTATATCTCATTTCATTCTCACAACTTATA
AGTGAATGAGTAAACTGAGACTTAAGGGAACTGAATCACTTAAATGTCACCTGGCTAACTGATGGCAGAGCCAGAGC
TTGAATTCATGTTGGTCTGACATCAAGGTCTTTGGTCTTCTCCCTACACCAAGTTACCTACAAGAACAATGACACCA
CACTCTGCCTGAAGGCTCACACCTCATACCAGCATACGCTCACCTTACAGGGAAATGGGTTTATCCAGGATCATGAG
ACATTAGGGTAGATGAAAGGAGAGCTTTGCAGATAACAAAATAGCCTATCCTTAATAAATCCTCCACTCTCTGGAAG
GAGACTGAGGGGCTTTGTAAAACATTAGTCAGTTGCTCATTTTTATGGGATTGCTTAGCTGGGCTGTAAAGATGAAG
GCATCAAATAAACTCAAAGTATTTTTAAATTTTTTTGATAATAGAGAAACTTCGCTAACCAACTGTTCTTTCTTGAG
TGTATAGCCCCATCTTGTGGTAACTTGCTGCTTCTGCACTTCATATCCATATTTCCTATTGTTCACTTTATTCTGTA
GAGCAGCCTGCCAAGAATTTTATTTCTGCTGTTTTTTTGCTGCTAAAGAAAGGGAACTAAGTCAGGATGTTAACAGA
AAAGTCCACATAACCCTAGAATTCTTAGTCAAGGAATAATTCAAGTCAGCCTAGAGACCATGTTGACTTTCCTCATG
TGTTTCCTTATGACTCAGTAAGTTGGCAAGGTCCTGACTTTAGTCTTAATAAAACATTGAATTGTAGTAAAGGTTTT
TGCAATAAAAACTTACTTTGG

FIGURE 22

CAGAGGTAGCCTGAAAGAAGCAGGAACTCCAGGATCCCAAACCAGAGCAGACCCTATAGTAAAGTATTTTTACATCT
TTTCCTTTCCCCAGAAGAGATCCCTAACCTATTGTTTTATTGACAGCCTTGCTGTTAGAGGCTCTTTCCCAGAAGTT
GGACGAAGAGGCTCAGGCGTTGCTGTTTCTTGTCTTCCAAGTCAAGTGGTTACTCTGGTAATGGATTGCCTCTCTCC
GAGCTTTCACCCTGGTGAGACTGTCCAGATCTAGTCTGTAAACCCAGCTTAGAAGCACTGTTGTAAAA**ATG**ACTGAA
GAGCCCATCAAGGAGATCCTGGGAGCCCCAAAGGCTCACATGGCAGCGACGATGGAGAAGAGCCCCAAGAGTGAAGT
TGTGATCACCACAGTCCCTCTGGTCAGTGAGATTCAGTTGATGGCTGCTACAGGGGGTACCGAGCTCTCCTGCTACC
GCTGCATCATCCCCTTTGCTGTGGTTGTCTTCATCGCCGGCATCGTGGTCACCGCGGTGGCTTACAGCTTCAATTCC
CATGGGTCTATTATCTCCATCTTTGGCCTGGTTGTTCTGTCATCTGGACTTTTTTTACTAGCCTCCAGTGCCTTGTG
CTGGAAAGTGAGACAAAGGAGCAAGAAAGCCAAGAGACGGGAGAGTCAAACAGCTCTCGTGGCAAATCAGAGAAGCT
TGTTTGCT**TGA**GACTGAATACGACCAAATGGGCCATTGGGCCTGGAAAACGTGCTCTGACTTTGTCACCCAATTCAC
CCAGAACCATGGTGGGAGAGAACAGACTTGGCGTTGGAGCAGACTGGAAGAATGGGGGTGGGAGGGTGGAGGGGCTT
CTCCTTTGTGAGGAATGACTCATGTCTTCTTTAACGACAAACTTAACCCTAAGGGCTACTTCTGAGACTGAAAAATC
AGCTTTCTATTTACATGAAACACTTTGGGGGTCATGGGAGTGCACAGCATTAGACAGTATTTGGTTCACCCTGTAAA
GTAGCCAAGAAAAGATGAGAAAAATCAAGATAGGCCTGGCACACTAGACATTTGCCTCCAAAAGAAATAACCTACAG
TCTTAAGATGTATCATAAAAATGTTCTGCCAAGGATCTAAATTACCTTGGGTTTCGCATATGTCTATGAAATTCTGT
GATAATTTTTTTCAATACATTGATTCACTGGCGTCTGTTTTCATTTTATACTTTTAATAACTCATCACTGGTGGTAC
TTTATCTTGAAAAGTAATATTTTTTATATTTTAACATTGGACAGTGTTAGCCAGTTGTAATGATGTATCAGAAGTAA
AGAAAAACCCATTAAAGTTATAGCTAATAGATGCTGTTGGGGGTTAAATTAATAGTAAAATAATCCAATATAGCACT
TTTGATGATTTTTATATAAAAGTCAACTGTACATTTCATTCAGAATAATAAATACTTATTGCTGCTAAAACTTCTTA
AATGGTTGTTTCTGCTATAGTTATTTCTATTGCAGTTCCAAATTGCCATCTTCCCTTGTCTCATTTGCAAGTTCTCA
ATTGTATTTCTCTCAAATGGACAGGTTCCTTCTTTACTGGAGGATTTTTGTTTTTATCATATTGGTTTTTCATTACT
TCTGAATAGTCTTAATTACGTTTACTAAATTCTAAAGGATTTCTGTGCTATTATAATTAGGAAATCAACGTCTTTGG
TCAGGAACTTTATAATGTGCTATTAAATGTATATTACATTTTTGTGG

FIGURE 23

ATGCTGTCACTGCTCCACGCATCAACGCTGGCAGTCCTTGGGGCTCTGTGTGTATATGGTGCAGGTCACCTAGAGCA
ACCTCAAATTTCCAGTACTAAAACGCTGTCAAAAACAGCCCGCCTGGAATGTGTGGTGTCTGGAATAACAATTTCTG
CAACATCTGTATATTGGTATCGAGAGAGACCTGGTGAAGTCATACAGTTCCTGGTGTCCATTTCATATGACGGCACT
GTCAGAAAGGAATCCGGCATTCCGTCAGGCAAATTTGAGGTGGATAGGATACCTGAAACGTCTACATCCACTCTCAC
CATTCACAATGTAGAGAAACAGGACATAGCTACCTACTACTGTGCCTTGTGGGAGGTGCCGGCTAGCCAACCAAGAGT
TGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTTTCC
CCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGA
GAAATTTTTCCCTGATGTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGA
ACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACAC
AGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATAAAGACAGATGT
CATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTG ·
CATATTACACGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTTAGAAGA
ACGGCTTTCTGCTGCAATGGAGAGAAATCA**TAA**CAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTG
TCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACT
ATTCTATCATTATTGTATAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT
AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCCAAATAGCGCCTGCTATA
GTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTGTTCTTTAATCAGATAACTGCCTGGAAGCCTTTCA
TTTTACACGCCCTGAAGCAGTCTTCTTTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTT
AAAAAAATG

FIGURE 24

```
GGGACAGGGCTGAGGATGAGGAGAACCCTGGGGACCCAGAAGACCGTGCCTTGCCCGGAAGTCCTGCCTGTAGGCCT
GAAGGACTTGCCCTAACAGAGCCTCAACAACTACCTGGTGATTCCTACTTCAGCCCCTTGGTGTGAGCAGCTTCTCA
ACATGAACTACAGCCTCCACTTGGCCTTCGTGTGTCTGAGTCTCTTCACTGAGAGGATGTGCATCCAGGGGAGTCAG
TTCAACGTCGAGGTCGGCAGAAGTGACAAGCTTTCCCTGCCTGGCTTTGAGAACCTCACAGCAGGATATAACAAATT
TCTCAGGCCCAATTTTGGTGGAGAACCCGTACAGATAGCGCTGACTCTGGACATTGCAAGTATCTCTAGCATTTCAG
AGAGTAACATGGACTACACAGCCACCATATACCTCCGACAGCGCTGGATGGACCAGCGGCTGGTGTTTGAAGGCAAC
AAGAGCTTCACTCTGGATGCCCGCCTCGTGGAGTTCCTCTGGGTGCCAGATACTTACATTGTGGAGTCCAAGAAGTC
CTTCCTCCATGAAGTCACTGTGGGAAACAGGCTCATCCGCCTCTTCTCCAATGGCACGGTCCTGTATGCCCTCAGAA
TCACGACAACTGTTGCATGTAACATGGATCTGTCTAAATACCCCATGGACACACAGACATGCAAGTTGCAGCTGGAA
AGCTGGGGCTATGATGGAAATGATGTGGAGTTCACCTGGCTGAGAGGGAACGACTCTGTGCGTGGACTGGAACACCT
GCGGCTTGCTCAGTACACCATAGAGCGGTATTTCACCTTAGTCACCAGATCGCAGCAGGAGACAGGAAATTACACTA
GATTGGTCTTACAGTTTGAGCTTCGGAGGAATGTTCTGTATTTCATTTTGGAAACCTACGTTCCTTCCACTTTCCTG
GTGGTGTTGTCCTGGGTTTCATTTTGGATCTCTCTCGATTCAGTCCCTGCAAGAACCTGCATTGGAGTGACGACCGT
GTTATCAATGACCACACTGATGATCGGGTCCCGCACTTCTCTTCCCAACACCAACTGCTTCATCAAGGCCATCGATG
TGTACCTGGGGATCTGCTTTAGCTTTGTGTTTGGGGCCTTGCTAGAATATGCAGTTGCTCACTACAGTTCCTTACAG
CAGATGGCAGCCAAAGATAGGGGGACAACAAAGGAAGTAGAAGAAGTCAGTATTACTAATATCATCAACAGCTCCAT
CTCCAGCTTTAAACGGAAGATCAGCTTTGCCAGCATTGAAATTTCCAGCGACAACGTTGACTACAGTGACTTGACAA
TGAAAACCAGCGACAAGTCAAGTTTGTCTTCCGAGAAAAGATGGGCAGGATTGTTGATTATTTCACAATTCAAAAC
CCCAGTAATGTTGATCACTATTCCAAACTACTGTTTCCTTTGATTTTTATGCTAGCCAATGTATTTTTACTGGGCATA
CTACATGTATTTTTGACATCAATGTTAAATTTCTTGCATGCCATAGGTCTTCAACAGGACAAGATAATGATGTAAATG
GTATTTTAGGCCAAGTGTGCACCCACATCCAATGGTGCTACAAGTGACTGAAATAATATTTGAGTCTTTCTGCTCAA
AGAATGAAGCTCCAACCATTGTTCTAAGCTGTGTAGAAGTCCTAGCATTATAGGATCTTGTAATAGAAACATCAGTC
CATTCCTCTTTCATCTTAATCAAGGACATTCCCATGGAGCCCAAGATTACAAATGTACTCAGGGCTGTTTATTCGGT
GGCTCCCTGGTTTGCATTTACCTCATATAAAGAATGGGAAGGAGACCATTGGGTAACCCTCAAGTGTCAGAAGTTGT
TTCTAAAGTAACTATACATGTTTTTACTAAATCTCTGCAGTGCTTATAAAATACATTGTTGCCTATTTAGGGAGTA
ACATTTTCTAGTTTTTGTTTCTGGTTAAAATGAAATATGGGCTTATGTCAATTCATTGGAAGTCAATGCACTAACTC
AATACCAAGATGAGTTTTTAAATAATGAATATTATTTAATACCACAACAGAATTATCCCCAATTTCCAATAAGTCCT
ATCATTGAAAATTCAAATATAAGTGAAGAAAAAATTAGTAGATCAACAATCTAAACAAATCCCTCGGTTCTAAGATA
CAATGGATTCCCCATACTGGAAGGACTCTGAGGCTTTATTCCCCCACTATGCATATCTTATCATTTTATTATTATAC
ACACATCCATCCTAAACTATACTAAAGCCCTTTTCCCATGCATGGATGGAAATGGAAGATTTTTTTGTAACTTGTTC
TAGAAGTCTTAATATGGGCTGTTGCCATGAAGGCTTGCAGAATTGAGTCCATTTCTAGCTGCCTTTATTCACATAG
TGATGGGGTACTAAAAGTACTGGGTTGACTCAGAGAGTCGCTGTCATTCTGTCATTGCTGCTACTCTAACACTGAGC
AACACTCTCCCAGTGGCAGATCCCCCTGTATCATTCCAAGAGGAGCATTCATCCCCTTTGCTCTAATGATCAGGAATGA
TGCTTATTAGAAAACAAACTGCTTGACCCAGGAACAAGTGGCTTAGCTTAAGTAAACTTGGCTTTGCTCAGATCCCT
GATCCTTCCAGCTGGTCTGCTCTGAGTGGCTTATCCCGCATGAGCAGGAGCGTGCTGGCCCTGAGTACTGAACTTTC
TGAGTAACAATGAGACACGTTACAGAACCTATGTTCAGGTTGCGGGTGAGCTGCCCTCTCCAAATCCAGCCAGAGAT
GCACATTCCTCGGCCAGTCTCAGCCAACAGTACCAAAAGTGATTTTTGAGTGTGCCAGGGTAAAGGCTTCCAGTTCA
GCCTCAGTTATTTTAGACAATCTCGCCATCTTTAATTTCTTAGCTTCCTGTTCTAATAAATGCACGGCTTTACCTTT
CCTGTCAGAAATAAACCAAGGCTCTAAAAGATGATTTCCCTTCTGTAACTCCCTAGAGCCACAGGTTCTCATTCCTT
TTCCCATTATACTTCTCACAATTCAGTTTCTATGAGTTTGATCACCTGATTTTTTTAACAAAATATTTCTAACGGGA
ATGGGTGGGAGTGCTGGTGAAAAGAGATGAAATGTGGTTGTATGAGCCAATCATATTTGTGATTTTTTAAAAAAAGT
TTAAAAGGAAATATCTGTTCTGAAACCCCACTTAAGCATTGTTTTTATATAAAAACAATGATAAAGATGTGAACTGT
GAAATAAATATACCATATTAGCTACCCACCAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 25

```
TGCTGTCAGGGCGCTCAAGCTGCCGGGGAAGAAGAGCCCAGACCTAGGGGAGTATGATCCACTTACCCAGGCTGACAGTGATGAGA
GCGAAGACGATCTGGTGCTTAACCTGCAGAAGAATGGAGGGGTCAAAAATGGGAAGAGTCCTTTGGGAGAAGCGCCAGAACCCGAC
TCAGATGCTGAGGTTGCAGAGGCTGCAAAGCCACATCTTTCAGAAGTCACCACGGAGGGCTACCCCTCAGAACCCCTTGGGGGCCT
GGAACAGAAGGCGGCCTCCTCCCTGGTGTCATATGTGCGCACGTCTGTCTTCCTGCTGACTTTGGGATCTCGATGATCCTGGTGC
TCCTGTGTGCTTTCCTGATCCCCTGTCCTCCCAGAGATCTGCACAGCACCTGGAGCCGCCACTTGGGCTCCCAGGGAGGTGGGGAC
CTGTCTCCAHTTGGAHTHGGCTGATGTGAHTCGGAGTAGCCCTGCGGTGAHGTGCHTCHTCTCCTHTTGTGAHTGTCAAGGAACGGGAGTGC
AGTAGGTGTCTCAAGACCAGCTGCTAAHTCTTGTGHTGGCCTTTCGGGGATGAAHTGGCAGCACACHTGHTGGHTCHTAGHTCHTHTCHTCCCHTGAGG
AGGCHTCGAGATATCACATGHTHTTGGAGCHTGATGCCAGGAAGCTHTGGCHTGAAHACCATCHTGCCGHTTGHTCACHGCGGACACAACAAGAHTGCHTC
AGCGCAHTTCAAHTGCAACHGTCAGGGAAAGCCAHTHTTGGACHTHTHTAAACCCAAACHTACHTHTGHTCCAACGGHTACCHTHTGGCHTGCCCCCAGHTHTGT
GGHTACHTGCCAGACHTHTGGAHTGAAGACGGHTGHTHTCGAGACCHTHTGHTGGHTHTCHTGGCCAHTHTGGGGGAAHTHTGCAGCCAGAHTCHTGHTGCHTHTHTCHTGC
TGGHTGHTCHTGGCCGGACCGGAAAHTCCAGHTGGGHTCGACCHTGHTGAAGHTACAACAHTCGHTHTGGAGHTHTGGGAAHTCHTGAHTHTGGHTCCHTCAGGHTHT
HTACAHTCACCACAAAHTGGGGCHTGHTCHTACAHTCCHTGHTHTHTGGCHTHTHTGGAAAHTAHTACAAGCHTGHTCGCACHTGCGGGACAHTHTHTHTGHTHTCAGGC
CCAAAHTCGAGACAGCHTCACCACCHTHTCHTCHTGCAGAHTAGAAGAGCCAGAAHTGGGAAAAGCGAAGAHTCCAHTCAACCHTGHTCHTGAGCHTCA
HTHTGAHTGHTHTHTACAGHTGAHTGGHTGHTHTGAACHTACHTCCAGAHTGGHTGAAGGCACCCAGAHTHTCCAACHTGCAGCAACCHTHTCHTGAHTHTACAACCAGA
CAAAGCCHTHTGHTGCHTGCHTHTCGGGGGCAAAAHTCHTGACACCHTHTACHTGGGCAHTHTGAGACHTHTCAAGGCCHTGCGCAGCCAGCCHTACHTCCHTGG
AHTAHTHTCACHTGAHTGAHTCAGACAHTHTAGAHTHTHTCCHTHTCHTGCAGAHTACAGGAHTGGAGHTHTGGGAHTGAAAAAGAHTGAHTGGHTHTGHTGGAHTGGHTG
ACHTCHTGGCHTCCAHTHTGHTHTGGAGHTHTACCGHTGCHTCCGHTGHTCACAHTGAAAGAAGAGCGCCAGCCACCHTCAGCAGHTHTACHTHTCAGACCAGAAG
HTCHTGHTCHTHTCCHTCHTHTCHTGGGCCGAAGGGCHTGHTCAGCHTGCAHTCHTCCCAAHTHTCCGAHTAHTCAHTCCHTAGGAACHTGAGCCGCCCAGCCHTHTCA
CCACCHTHTHTACCHTCCHTGCAHTCCHTGCGHTCCCCHTCCAHTCCHTHTCHTGGAHTCHTGGCCAAACACCACCGGCACAHTGACGGCHTHTCAGAGGHTHTG
GAAHTHTAACGACCHTCHTGGAAAGAHTGCCHTHTHTHTAHTGHTHTACCAGGACAACAGGGCCAAGCHTCCGAAGGCCAHTCCAGCAGCCCHTGGHTGGHTC
AGCAAGCHTHTAGHTCHTACGGHTGGGCACHTAAHTGGAGGGGCCAGAHTGGCHTCAGCGAGHTGCCACCCCCAAAAHTHTGGCCGHTGGGGAGCHT
GCGAAGAHTHTHTCHTCHTCHTAGGAHTAAAGHTHTHTGHTHTGAAGCHTCCCHTACGAGAHTCTAATCHTGAHTGGAAHTCHTHTCAGHTGCAGAAGAAGHTGAAC
AGAGHTGGAHTACCCHTCHTCHTACHTCHTCCHTGHTCACHTGHTAAAAHTCAGHTHTCHTAHTGGAGGAAGACHTHTCHTHTCGHTCCHTCAHTHTHTACCACCHTCCCHT
GAHTGGHTHTGCAAAGGCHTHTGGGAAGGCAHTGHTHTGGAGHTCHTHTHTGACGGCAGCAHTGAHTCHTAHTHTHTGGCHTGGGGCAHTCHTHTACCHTACCHTHTHTHTCA
GHTCCCHTGCAHTHTAAHTCCCCHTCHTAGGAACHTCHTGCCHTGGAHTCGHTHTGCGGGAAHTGHTCAAHTCHTCHTHTAAGHTAHTHTHTAAHTHTHTHTHTGGHTAHTGHTCHT
AAHTHTHTAHTGAAGHTCHTHTGCHTGGGAHTGCCHTACCHTCHTGAHTGHTGAGGGCCACHTHTAHTHTHTCHTCHTCCCHTAHTHTCHTHTHTCCCACCHTGHTACCHTHTGG
HTAGHTGHTHTGHTGGHTGAAGHTHTAHTHTHTHTCCAGGHTAHTGHTCCHTAAGCHTHTCAGGGAHTCCAGHTHTCHTGHTGHTCCHTHTCHTGCHTGAAHTAHTAHTCHTGCHTHTGCHTHTHTGHTHT
HTGGHTCAHTHTHTGAGACHTHTCCAGAHTGCCCHTACCHTCHTGAHTGHTHTGAGGGCCACHTHTAHTHTHTCHTCHTCCCHTAHTHTCHTHTHTCCCACCHTGHTACCHTHTGG
CHTACHTHTCCAAAHTHTGHTAGACAGAAHTGAGAAAGAHTHTHTAHTAGHTGGAAGACHTGAGHTHTAGCCAHTCCAAGCAHTHTHTCAHTCHTCHTCHTGHTHTHTA
HTAHTCCHTAHTHTHTCCHTHTAGAHTHTHTHTCCAHTCAHTGHTCHTAHTHTHTAAGHTGACCACAAGAAHTAACHTAHTAHTHTCCHTAHTCACAAGGGGACAAGAGGAHT
GHTAGHTCHTCAGHTGACCCAHTCHTCHTGACCAAGHTCCACAHTGHTGHTGHTHTAHTAHTGHTGGCHTCHTGAHTGGHTHTCHTGCCAGHTCAHTGAHTCHTHTHTHTHTHTCHT
GHTGGCGACAHTCAGAAGHTGHTAHTGHTHTHTGCAHTGCHTGHTCHTHTCAACHTHTAGAGGAGAACHTGGAAGHTCAGGAGCCHTHTHTGAHTGHTCCHTHTAHTCCHTG
CHTGHTAHTGHTCHTHTCHTCHTGCHTACHTHTHTHTHTHTCHTAHTAGGGCACCCHTCCHTHTAGCHTCCCCHTCACHTCHTGHTHTHTHTCHTHTCHTCHTAAHTHTCAGGGAHTAHTGHTHTHT
CHTGGACHTHTHTHTHTCHTHTCHTGCHTACHTHTGAGHTCGHTGGAHTCCHTGGHTHTHTGHTCHTAHTCHTHTHTGHTGCHTAGHTCHTHTGHTCHTHTGHTAGAGCCHTHTGAAGCAHTHTG
HTAHTHTHTHTGGGAAAAHTHTCHTHTCHTGHTAAAHTACHTAHTAACHTHTHTHTAHTAAAHTGGHTHTAAGHTHTAHTHTHTAGAAHTAHTCHTCCAGHTGCHTHTACHTHTCHTCCCHT
HTCHTHTCHTGHTAHTAAAHTCHTGCHTACHTHTCAAHTHTAAGHTHTCHTCCHTCHTAAACHTHTHTAGGHTCAHTHTGHTHTHTAHTAHTACAGAAAAHTHTCAAHTGHTHTAGCG
GAHTGGAAAACHTGCHTHTCHTHTGAAHTAACCHTGAHTAGGHTCAHTCCCHTGAGHTGCACCHTCAGGHTHTCHTCHTCHTHTHTACCHTGGGCHTGHTAHTCHTHTHTHTHT
HTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTGAGACAGAGHTHTHTGCHTCHTHTGHTCGCCCAGGCHTGGAGHTGCAGHTGGCACAAHTCHTCGGCHTCACHTGCAAC
CHTCGCCHTCCHTGGGHTHTCAAGCGAHTHTCHTCCAGCCHTHTAGCCHTCCCAAGHTAGCHTGGGACHTACAGGHTGCCCGCHTACCAHTGCCHTGGCHTAAHT
HTHTHTHTHTHTHTGHTAHTHTHTHTHTAGHTAGAGACGGGGHTHTHTCACCAHTGHTHTGGCCAGGCHTGGHTCHTCGAACHTCCHTGACCHTCAGAHTAAHTCCACCHTG
CHTHTCHTGCCHTCCCAAAGHTGCHTGGGAHTHTACAAGGCGHTGAGCCACCAHTGCCCGGCHTGGGCHTHTGHTAHTCHTHTHTHTAGCHTHTGHTGHTHTAGHTAAAAGG
AHTHTCHTAGAAAAHTHTAHTGAAGHTCCAGAHTHTCAAAGGGAHTCHTCHTGHTHTAAHTHTACCCACHTGACAGGCAHTHTAHTGACCHTAACAGGAGGHTHTGGHTA
GCAGHTAGAHTCCAAGCAHTGCAHTGHTHTGCCHTGGCCHTGHTAGAHTHTGGCCHTHTAHTCAGGHTHTHTCHTGGGHTGCCHTCHTGCCHTHTAAGAHTCCHTGAAGGC
AAAHTHTHTHTGHTHTHTCAACAGHTHTHTGGAAGHTCAHTCHTGHTGGGHTCCAGCHTHTGACHTHTHTGGAGGAAHTAAGAAGAHTACHTHTCHTAGAGHTAHTGGGAAHTG
AHTHTCCAGAHTAAHTHTHTCHTGGGAHTHTHTGAAHTCHTACHTHTGAGHTHTHTAAGGGCCHTGGGACCHTAAHTHTHTGGHTHTHTAGHTAHTAGAAHTHTHTGAAGAAHTHTAA
HTHTHTAHTAGGCAGCHTGAAHTACCCAAAACHTHTGGGHTGGHTGGHTCCHTGHTGGHTHTHTGGCHTGAGCHTGHTCCGGGCAHTAACCHTGGHTHTCHTCHTGHTHTAHTG
HTHTAAGGCHTHTHTCHTGGGAAGCCAGCCACHTCHTGCGCAGGAGHTGAAACAHTGAAGHTHTGHTHTHTHTCHTGAGGACCHTGHTHTHTGGHTGGGAHTHTGHTHTHTG
GGCAGAGGACHTGHTGHTHTHTAHTGCAGGGCAAAHTCCCAGAAAGAHTAAGAGGAAGCHTAGAGAAACHTHTAAHTGHTACCHTGAAHTHTCHTHTCAHTGGHTG
HTAHTHTHTGCAAACHTAACHTHTAACAHTAGAHTHTCHTHTHTHTGACHTAHTGAHTGGHTAAGHTHTHTGAAHTCHTCHTCCHTHTGCCAAACAACAHTHTAHTAAGHTHTAGHTHTHTHT
CHTHTCHTHTCCHTCHTHTGCAGCCGGHTACAGAAAGGHTGHTAAGHTGGHTGGCHTGAAAAHTHTGAGGAAGCHTHTCAHTCHTGACCAAHTGHTGGGHTGCHTGGHTHT
HTCHTHTGHTGAAAHTGHTGHTCCCHTAAGGCHTCCHTHTCCHTHTGCAGGCAGCCACCCACCCACGGHTGHTCHTAAGAHTAGGACAHTGCHTCCHTHTHTCHTHTHTC
HTCHTAHTCCCAHTCCHTGAGGHTHTGCCGGCAAAGCCAAHTAHTGACCACHTACHTGAGAAAHTAGHTAAHTGACHTHTCHTACAAAHTGCAAGGGHTCHTHTAC
CCHTCCHTCHTHTHTCCCHTHTAAACACCCHTCCCHTHTHTHTCCHTHTAGACCCCGHTHTHTHTGCCAHTCCCCCAAAHTGHTGHTGGHTAHTGGHTGAAACHTAAHTCCC
CHGAAHTGHTGAAHTHTGCHTAHTCCHTHTAHTHTHTGCCCHTAHTHTAAAGAAGAGCCAGCHTGGHTAHTAHTHTGHTCAGGAAGCACHTAHTHTHTAAAAHTGHTGAACHTG
HTHTAHTAGAGHTAAAHTAAAHTAAAHTACHTCHTACAGG
```

FIGURE 26

```
CGCTTAGAACTGTGTTGAGCTCTCACCCATCACGATGAGCAACAAATTCCTGGGCACCTGGAAACTTGTCTCTAGCG
AGAACTTTGACGATTACATGAAAGCTCTGGGTGTGGGGTTAGCCACCAGAAAACTGGGAAATTTGGCCAAACCCACT
GTGATCATCAGCAAGAAAGGAGATATTATAACTATACGAACTGAAAGTACCTTTAAAAATACAGAAATCTCCTTCAA
GCTAGGCCAGGAATTTGAAGAAACCACAGCTGACAATAGAAAGACCAAGAGCATCGTAACCCTGCAGAGAGGATCAC
TGAATCAAGTGCAGAGATGGGATGGCAAAGAGACAACCATAAAGAGAAAGCTAGTGAATGGGAAAATGGTAGCGGAA
TGTAAAATGAAGGGCGTGGTGTGCACCAGAATCTATGAGAAGGTCTGAAAATCATTTCTTCATTGAAGTGGCTTTT
TATCATTTAATGATGGAAATCAATTGCTTCCATTGACAAAACTGAATACACTGCAAATATTTGTTTTTGCTTTTGTC
TTAATATATCAGATATGCAAAGGCCTAAACTGAGAATTAATCTAAAAGTCAGTGTTATTTAAACATTTTCAATGTGC
ATGCATGTCATTATTACATCAAAGCATATATATTGGCCAGACACAAACAGTTGATGATGTCATTCAATTAACTACAA
AATTCTAATCTATGTTGAACTTTGTATACTTGAAATGATAATAAAAAGGATATAATTTCTTAGTAAAATGAAATCAA
AGTATTGATCAGGGTAGCAAACTCAAATGCTGACAGGGGCCAGAGGAGATATGGGGAAGGAGCATCAGAAATGAGGC
AAGCTAGGAGAATGGGCTATTATAATGTAAAGAATTGTAGTCTCAGTTAAAAGGGGTAGCCTCTACTCCAGCCAACA
TTTTAAAATTAATGGATAATTTATAGACAGTTAAATTTATAGACAGTTAAGTAAAAATGGATAATTTATAGACAGAT
AATTTATAGACAGGTAAATGTGAGTTAAATATAACTCACATCCCACTCAAGACACAAAACATTTTCTTAATCCTAGT
ACATTTTTTTCTGTCCCTTCCCAATCAGTGTCCTTTTCTGTTCCACCCCTACCAAAAGCAAGTAGTGGTTTGGTTTC
TATCATATAGATTAATTTTACCTGCTCATATGAAGGGAATTGTACATCATGCATTCTTTTCTGTTTGCCTTTTTTAA
ATTCAGCATCATGTTTTTGTGATACATCCACATTGTTGCATGCAGCTGTAGTTTGTTTCTTTTTATTACCAAGTACT
ATTTCATTGTATGAATATATCACAGTTTATCCATTCTACTATTAAGACAATTGAGCTATTTCTAATTTTCGGCTGCT
ATGAATAAAGCTGCTACAAATATTTTTGTACAAGACTTTTTGTAAACATAGGAGTCCCTTTATTTTTAAATAAATAAC
TAGTCATATCATTAGGTCCAGTAATTGTTGACAGGCAGGAACGGGGGACCATTGCATTGTGCCCAAGTAATAATAAA
ACTATTTCAGATGTATTATATGATTGAGCAAATGAGAAAACATGTTGATGTTGATGGGAGTCAGGATGTTCACTATG
GAAAAACAAATATACAAATATGAAATGAGGGAAGGCAAGAAAGAACCATGTGGAAATGGAATAGAATTGGTATAAAT
TCATAATTTCTAAACCATGTATATGTACGTTTATATGTATTATAATTGCATACACATGCCTCCATGCATATATGTGT
GTGATAATACACATGCATTTATGTGCGTGTGTGTATACACATGCATATATTTACTAATCCTATCTGCCAAAATGGCT
TAGACACAAAAACACCTCAGCAGAAATGAATATACCTAGCACTCAGATCTTCGTGTCTAATATAGTTTGCCACTAAA
AGGAACCAAGGCTACTTGGAAAAATGGATGATTCCAAAGCAAGGGCAAGGTAGGAACAAGATGAGCTTGAAATATCT
TGTTATGCCAGAAAGTAATGTTAAAAAAAAAAAATAGAGGTATATTGCCAAAACATAGAGCCAGCTTGAAGGGGCTC
CCACTGGCCAAACTTGAGCCAATCTGAGAGCAAAATAATTGAGAAAAATAAATAACAAGATAATTGAAAAAAAATAA
AATAAATTCCCAATTCCATTGGGAAAAAATAGAAATCCATGAATCCATACTGATATAATGATAGATAATCGATAGAT
AGGTAGATGACTGACAGATGATAGATAGATATATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGA
GAAATGAGAAGGTTTCTCTTTTGCAATAGAACGTCATGGACCAGTGTTAAATGTGAGTGGAAGGAGTTCTGGAATTG
GAAAACCATCATTTTTCAACCATCACAGTAAATATGGCTCAGGCAAGAATTATCAATCAATGCTAAAGCTAGGGGGA
AATTTCGCTTAGGAGCAGGATATTAGGGTATTAGTCTGGGCTTAAAGTATCTCCTCACAGATTGTTGTTAGTTTCTG
GGGAAAGAATAGTAACCATGCAATGGAAAAAAATGGACAACCTCTTGACTAGGTTATCAAAATTAACCTCACCAATA
AAGGGTGGATGTTCAACATGTGCCTTCAAATGTGACCCACTGAGAAGGAAACAACATCACTGTAACAACAACAACCA
GAAACGACAGGGGGTTTTGACTGAATTCTTCAAAAATGTCAATGTCATAGAAGACAAAGAAAGGTTGTGGAAATGTT
TCAGATTAAATGATAGTAAAAACACCTGACAACTAAACATAGTAAGTAATACTAGACTGGATTCTGTACCAGAGGTA
ACATAAGTGCTCCAAAGGACAATGTTAGGTCAACTGGCAAATTGGAATATAGACAGTCAATCAGATAAGAAGTATAC
TTTGATTAAGTAAAAAAAATCCCTATTCTTGGAAAATACACAATAAAGTATTTTGAGGTAAAAGGGCCCATAATGTATG
CAATCTACTCTCAAAAAATTCAGAAACATATATTTGTGTGCATTTGCATGTGCAACAGTACACACAAACATACATAA
AGAGAGCAATTGATAAGGCAAATAAGGTAACATTTAACAATAATCTGATACACATAAATAGAGAAAGAGCAATTGAT
AAAGTAAATGAGGTAAAATTTAACAATAATCTGAGCAAAAGGTATATGTGTTTCTTTGAGACAGTCTGATTCCTTGC
AACTTATTCTGTAAGTTGGAACTTATTTCCAAACATGATTGAAAAAAAACCCCGCACTTGGCAACTTCTTCTTCTTTT
TCAGCCTAGAAATGTCTGTGTTAAGTGGTTTTTTTATTTATTGTTGTTGTTTTGTTGGTTATTGTTGTTTTTGTTGCCAGG
CTCCAACTCACAAAATACGAGTTTAAAAACTGCGTTGTTATTTTTAGAGATTTGTGATAATACAACTTGTTATAAAA
TTTATTCCTCAATAAATATAATTTCTCTACTATGCAAAAAAAAAAAAAAA
```

FIGURE 27

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGGCTGGGCGCGCCCCCC
GGGCCCCGCCGTGGGC**ATG**GGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCG
CCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACC
CCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGC
GGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGA
TCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC
TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTA
CACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTG
TCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT
TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATCCCCAA
CGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGG
GTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAA
ATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGA
CAGTGGCACCACGCTGCTGCGCGCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGA
TTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTAC
TTCCCTAAAATCTCCATCTACCTGAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTA
CATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTGG
TGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC
CCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTG
TGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCC
TCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCGTGACCCTGAGGTCGTCAATGAT
GAGTCCTCTCTGGTCAGACATCGCTGGAAA**TGA**ATAGCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAG
AAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTT
CTGCTCCCAGATGCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAG
AAAAATAATTAAAAAAAAAACTTCATTCTAA

FIGURE 28

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGGCTCTGGCAGGCTCCT
GGCAGC**ATG**GCAGTGAAGCTTGGGACCCTCCTGCTGGCCCTTGCCCTGGGCCTGGCCCAGCCAGCCTCTGCCCGCCG
GAAGCTGCTGGTGTTTCTGCTGGATGGTTTTCGCTCAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTT
TCAAAGAGATTGTGAGCAGGGGAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTAT
TATACCCTAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACTACATGTGGGACCCCACCACCAACAA
GTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTGGAATGGATCAGAACCTCTGTGGGTCACTC
TGACCAAGGCCAAAAGGAAGGTCTACATGTACTACTGGCCAGGCTGTGAGGTTGAGATTCTGGGTGTCAGACCCACC
TACTGCCTAGAATATAAAAATGTCCCAACGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAA
GAGTGGCCGGGCCGACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGGGCCTGCATCTC
CGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGACCAAGTGGATCCAGGAGCGGGGCCTG
CAGGACCGCCTGAACGTCATTATTTTCTCGGATCACGGAATGACCGACATTTTCTGGATGGACAAAGTGATTGAGCT
GAATAAGTACATCAGCCTGAATGACCTGCAGCAAGTGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCTG
GGAAACACTCTGAGATATATAACAAACTGAGCACAGTGGAACACATGACTGTCTACGAGAAAGAAGCCATCCCAAGC
AGGTTCTATTACAAGAAAGGAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGATGAAGGCTGGTTCATAACTGAGAA
TCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCAGGCGGGAAGGTTGGCAGCGTGGATGGCACGGCTACGACA
ACGAGCTCATGGACATGCGGGGCATCTTCCTGGCCTTCGGACCTGATTTCAAATCCAACTTCAGAGCTGCTCCTATC
AGGTCGGTGGACGTCTACAATGTCATGTGCAATGTGGTGGGCATCACCCCGCTGCCCAACAACGGATCCTGGTCCAG
GGTGATGTGCATGCTGAAGGGCCGCGCCGGCACTGCCCCGCCTGTCTGGCCCAGCCACTGTGCCCTGGCACTGATTC
TTCTCTTCCTGCTTGCA**TAA**CTGATCATATTGCTTGTCTCAGAAAAAAACACCATCAGCAAAGTGGGCCTCCAAAGC
CAGATGATTTTCATTTTATGTGTGAATAATAGCTTCATTAACACAATCAAGACCATGCACATTGTAAATACATTATT
CTTGGATAATTCTATACATAAAAGTTCCTACTTGTTAAA

FIGURE 29

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACC**ATG**AAGTTCTTAGCAGTCCTGGTACTCTTGGGAGTTTC
CATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTGACACGTATCCAGCTACTGGTCCTGCTGATG
ATGAAGCCCCTGATGCTGAAACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCT
GCTTCTACCACTGCTCGTAAAGACATTCCAGTTTTACCCAAATGGGTTGGGGATCTCCCGAATGGTAGAGTGTGTCC
C**TGA**GATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACTATTCATGCTTCCTGTGATTTCATCCAACTACTTA
CCTTGCCTACGATATCCCCTTTATCTCTAATCAGTTTATTTTCTTTCAAATAAAAAATAACTATGAGCAACATAAAA
AAAAAAAAA

FIGURE 30

GACTACGGGGAGAGAGAGGAGACCAGGACAGCTGCTGAGACCTCTAAGAAGTCCAGATACTAAGAGCAAAG**ATG**TTT
CAAACTGGGGGCCTCATTGTCTTCTACGGGCTGTTAGCCCAGACCATGGCCCAGTTTGGAGGCCTGCCCGTGCCCCT
GGACCAGACCCTGCCCTTGAATGTGAATCCAGCCCTGCCCTTGAGTCCCACAGGTCTTGCAGGAAGCTTGACAAATG
CCCTCAGCAATGGCCTGCTGTCTGGGGGCCTGTTGGGCATTCTGGAAAACCTTCCGCTCCTGGACATCCTGAAGCCT
GGAGGAGGTACTTCTGGTGGCCTCCTTGGGGGACTGCTTGGAAAAGTGACGTCAGTGATTCCTGGCCTGAACAACAT
CATTGACATAAAGGTCACTGACCCCCCAGCTGCTGGAACTTGGCCTTGTGCAGAGCCCTGATGGCCACCGTCTCTATG
TCACCATCCCTCTCGGCATAAAGCTCCAAGTGAATACGCCCCTGGTCGGTGCAAGTCTGTTGAGGCTGGCTGTGAAG
CTGGACATCACTGCAGAAATCTTAGCTGTGAGAGATAAGCAGGAGAGGATCCACCTGGTCCTTGGTGACTGCACCCA
TTCCCCTGGAAGCCTGCAAATTTCTCTGCTTGATGGACTTGGCCCCCTCCCCATTCAAGGTCTTCTGGACAGCCTCA
CAGGGATCTTGAATAAAGTCCTGCCTGAGTTGGTTCAGGGCAACGTGTGCCCTCTGGTCAATGAGGTTCTCAGAGGC
TTGGACATCACCCTGGTGCATGACATTGTTAACATGCTGATCCACGGACTACAGTTTGTCATCAAGGTC**TAA**GCCTT
CCAGGAAGGGGCTGGCCTCTGCTGAGCTGCTTCCCAGTGCTCACAGATGGCTGGCCCATGTGCTGGAAGATGACACA
GTTGCCTTCTCTCCGAGGAACCTGCCCCCTCTCCTTTCCCACCAGGCGTGTGTAACATCCCATGTGCCTCACCTAAT
AAAATGGCTCTTCTTATGCA

FIGURE 31

```
GGGCGCTGGGAGACACCGGACGCCCGCTCGGCTGCGCTGCGGCTCAGGCCCCCGCTCGGGCCCGACCCGCTCGGTCA
CCGCCGGCTCGGGCGCGCACCTGCCGGCTGCGGCCCCAGGGCCATGCGGGAGGCCCACGAGGAGGCCGGCGGCCACGC
GCATCCCGTAGCCCAGGTGGCCCAGGTCTGCACCGCGGCGGCCTCGGCGCCATGGAGCCCCCGTATTCGCTGACGGC
GCACTACGATGAGTTCCAAGAGGTCAAGTACGTGAGCCGCTGCGGCGCGGGGGGCGCGCGCGGGGCCTCCCTGCCCC
CGGGCTTCCCGTTGGGCGCTGCGCGCAGCGTCACCGGGGCCCGGTCCGGGCTGCCGCGCTGGAACCGGCGCGAGGTG
TGCCTGCTGTCGGGGCTGGTGTTCGCCGCCGGCCTCTGCGCCATTCTGGCGGCTATGCTGGCCCTCAAGTACCTGGG
CCCGGTCGCGGCCGGCGGCGGCGCCTGTCCCGAGGGCTGCCCTGAGCGCAAGGCCTTCGCGCGCGCCGCTCGCTTCC
TGGCCGCCAACCTGGACGCCAGCATCGACCCATGCCAGGACTTCTACTCGTTCGCCTGCGGCGGTTGGCTGCGGCGC
CACGCCATCCCCGACGACAAGCTCACCTATGGCACCATCGCCGGCCATCGGCGAGCAAAACGAGGAGCGCCTACGGCG
CCTGCTGGCGCGGCCCGGGGGTGGGCCTGGCGGCGCGGCCCAGCGCAAGGTGCGCGCCTTCTTCCGCTCGTGCCTCG
ACATGCGCGAGATCGAGCGACTGGGCCCGCGACCCATGCTAGAGGTCATCGAGGACTGCGGGGGCTGGGACCTGGGC
GGCGCGGAGGAGCGTCCGGGGGTCGCGGCGCGATGGGACCTCAACCGGCTGCTGTACAAGGCGCGAGGGCGTGTACAG
CGCCGCCGCGCTCTTCTCGCTCACGGTCAGCCTGGACGACAGGAACTCCTCGCGCTACGTCATCGCATTGACCAGG
ATGGGCTCACCCTGCCAGAGAGGACCCTGTACCTCGCTCAGGATGAGGACAGTGAGAAGATCCTGGCAGCATACAGG
GTGTTCATGGAGCGAGTGCTCAGCCTCCTGGGTGCAGACGCTGTGGAACAGAAGGCCCAAGAGATCCTGCAAGTGGA
GCAGCAGCTGGCCAACATCACTGTGTCAGAGTATGACGACCTACGGCGAGATGTCAGCTCCATGTACAACAAGGTGA
CGCTGGGGCAGCTGCAGAAGATCACCCCCCACTTGCGGTGGAAGTGGCTGCTAGACCAGATCTTCCAGGAGGACTTC
TCAGAGGAAGAGGAGGTGGTGCTGCTGGCGACAGACTACATGCAGCAGGTGTCGCAGCTCATCCGCTCCACACCCCA
CCGGGTCCTGCACAACTACCTGGTGTGGCGCGTGGTGGTGGTGCCTGAGTGAACACCTGTCCCCGCCATTCCGTGAGG
CACTGCACGAGCTGGCACAGGAGATGGAGGGCAGCGACAAGCCACAGGAGCTGGCCCGGGTCTGCTTGGGCAGGCC
AATCGCCACTTTGGCATGGCGCTTGGCGCCCTCTTTGTACATGAGCACTTCTCAGCCGCCAGCAAAGCCAAGGTGCA
GCAGCTAGTGGAAGACATCAAGTACATCCTGGGCCAGCGCCTGGAGGAGCTGGACTGGATGGACGCCGAGACCAGGG
CTGCTGCTCGGGCCAAGCTCCAGTACATGATGGTGATGGTCGGCTACCCGGACTTCCTGCTGAAACCCGATGCTGTG
GACAAGGAGTATGAGTTTGAGGTCCATGAGAAGACCTACTTCAAGAACATCTTGAACAGCATCCCCTTCAGCATCCA
GCTCTCAGTTAAGAAGATTCGGCAGGAGGTGGACAAGTCCACGTGGCTGCTCCCCCCACAGGCGCTCAATGCCTACT
ATCTACCCAACAAGAACCAGATGGTGTTCCCCGCGGGCATCCTGCAGCCCACCCTGTACGACCCTGACTTCCCACAG
TCTCTCAACTACGGGGGCATCGGCACCATCATTGGACATGAGCTGACCCACGGCTACGACGACTGGGGGGGCCAGTA
TGACCGCTCAGGGAACCTGCTGCACTGGTGGACGGAGGCCTCCTACAGCCGCTTCCTGCGAAAGGCTGAGTGCATCG
TCCGTCTCTATGACAACTTCACTGTCTACAACCAGCGGGTGAACGGGAAACACACGCTTGGGGAGAACATCGCAGAT
ATGGGCGTCCTCAAGCTGGCCTACCACGCCTATCAGAAGTGGGTGCGGGAGCACGGCCCAGAGCACCCACTTCCCCG
GCTCAAGTACACACATGACCAGCTCTTCTTCATTGCCTTTGCCCAGAACTGGTGCATCAAGCGGCGGTCGCAGTCCA
TCTACCTGCAGGTGCTGACTGACAAGCATGCCCCTGAGCACTACAGGGTGCTGGGCAGTGTGTCCCAGTTTGAGGAG
TTTGGCCGGGCTTTCCACTGTCCCAAGGACTCACCCATGAACCCTGCCCACAAGTGTTCCGTGTGGTGAGCCTGGCT
GCCCGCCTGCACGCCCCCACTGCCCCCGCACGAATCACCTCCTGCTGGCTACCGGGGCAGGCATGCACCCGGTGCCA
GCCCCGCTCTGGGCACCACCTGCCTTCCAGCCCCTCCAGGACCCGGTCCCCCCTGCTGCCCCTCACTTCAGGAGGGGC
CTGGAGCAGGGTGAGGCTGGACTTTGGGGGGCTGTGAGGGAAATATACTGGGGTCCCCAGATTCTGCTCTAAGGGGG
CCAGACCCTCTGCCAGGCTGGATTGTACGGGCCCCACCTTCGCTGTGTTCTTGCTGCAAAGTCTGGTCAATAAATCA
CTGCACTGTTAAAAAAAAAA
```

FIGURE 32

GAGAGAACAGCGTGAGCCTGTGTGCTTGTGTGCTGAGCCCTCATCCCCTCCTGGGGCCAGGCTTGGGTTTCACCTGC
AGAATCGCTTGTGCTGGGCTGCCTGGGCTGTCCTCAGTGGCACCTGCATGAAGCCGTTCTGGCTGCCAGAGCTGGAC
AGCCCCAGGAAAACCCACCTCTCTGCAGAGCTTGCCCAGCTGTCCCCGGGAAGCCAAATGCCTCTCATGTAAGTCTT
CTGCTCGACGGGGTGTCTCCTAAACCCTCACTCTTCAGCCTCTGTTTGACCATGAAATGAAGTGACTGAGCTCTATT
CTGTACCTGCCACTCTATTTCTGGGGTGACTTTTGTCAGCTGCCCAGAATCTCCAAGCCAGGCTGGTTCTCTGCATC
CTTTCAATGACCTGTTTTCTTCTGTAACCACAGGTTCGGTGGTGAGAGGAAGCCTCGCAGAATCCAGCAGAATCCTC
ACAGAATCCAGCAGCAGCTCTGCTGGGGACATGGTCCATGGTGCAACCCACAGCAAAGCCCTGACCTGACCTCCTGA
TGCTCAGGAGAAGCC**ATG**GGCCCCTCCTGTCCTGTGTTCCTGTCCTTCACAAAGCTCAGCCTGTGGTGGCTCCTTCT
GACCCCAGCAGGTGGAGAGGAAGCTAAGCGCCCACCTCCCAGGGCTCCTGGAGACCCACTCTCCTCTCCCAGTCCCA
CGGCATTGCCGCAGGGAGGCTCGCATACCGAGACTGAGGACCGGCTCTTCAAACACCTCTTCCGGGGCTACAACCGC
TGGGCGCGCCCGGTGCCCAACACTTCAGACGTGGTGATTGTGCGCTTTGGACTGTCCATCGCTCAGCTCATCGATGT
GGATGAGAAGAACCAAATGATGACCACCAACGTCTGGCTAAAACAGGAGTGGAGCGACTACAAACTGCGCTGGAACC
CCGCTGATTTTGGCAACATCACATCTCTCAGGGTCCCTTCTGAGATGATCTGGATCCCCGACATTGTTCTCTACAAC
AATGCAGATGGGGAGTTTGCAGTGACCCACATGACCAAGGCCCACCTCTTCTCCACGGGCACTGTGCACTGGGTGCC
CCCGGCCATCTACAAGAGCTCCTGCAGCATCGACGTCACCTTCTTCCCCTTCGACCAGCAGAACTGCAAGATGAAGT
TTGGCTCCTGGACTTATGACAAGGCCAAGATCGACCTGGAGCAGATGGAGCAGACTGTGGACCTGAAGGACTACTGG
GAGAGCGGCGAGTGGGCCATCGTCAATGCCCACGGGCACCTACAACAGCAAGAAGTACGACTGCTGCGCCGAGATCTA
CCCCGACGTCACCTACGCCTTCGTCATCCGGCGGCTGCCGCTCTTCTACACCATCAACCTCATCATCCCCTGCCTGC
TCATCTCCTGCCTCACTGTGCTGGTCTTCTACCTGCCCTCCGACTGCGGCGAGAAGATCACGCGTGTGCATTCGGTG
CTGCTGTCACTCACCGTCTTCCTGCTGCTCATCACTGAGATCATCCCGTCCACCTCGCTGGTCATCCCGCTCATCGG
CGAGTACCTGCTGTTCACCATGATCTTCGTCACCCTGTCCATCGTCATCACCGTCTTCGTGCTCAATGTGCACCACC
GCTCCCCCAGCACCCACACCATGCCCCACTGGGTGCGGGGGGCCCTTCTGGGCTGTGTGCCCCGGTGGCTTCTGATG
AACCGGCCCCCACCACCCGTGGAGCTCTGCCACCCCCTACGCCTGAAGCTCAGCCCCTCTTATCACTGGCTGGAGAG
CAACGTGGATGCCGAGGAGAGGGAGGTGGTGGTGGAGGAGGAGGACAGATGGGCATGTGCAGGTCATGTGGCCCCCT
CTGTGGGCACCCTCTGCAGCCACGGCCACCTGCACTCTGGGGCCTCAGGTCCCAAGGCTGAGGCTCTGCTGCAGGAG
GGTGAGCTGCTGCTATCACCCCACATGCAGAAGGCACTGGAAGGTGTGCACTACATTGCCGACCACCTGCGGTCTGA
GGATGCTGACTCTTCGGTGAAGGAGGACTGGAAGTATGTTGCCATGGTCATCGACAGGATCTTCCTCTGGCTGTTTA
TCATCGTCTGCTTCCTGGGGACCATCGGCCTCTTTCTGCCTCCGTTCCTAGCTGGAATGATC**TGA**CTGCACCTCCCT
CGAGCTGGCTCCCAGGGCAAAGGGGAGGGTTCTTGGATGTGGAAGGGCTTTGAACAATGTTTAGATTTGGAGATGAG
CCCAAAGTGCCAGGGAGAACAGCCAGGTGAGGTGGGAGGTTGGAGAGCCAGGTGAGGTCTCTCTAAGTCAGGCTGGG
GTTGAAGTTTGGAGTCTGTCCGAGTTTGCAGGGTGCTGAGCTGTATGGTCCAGCAGGGGAGTAATAAGGGCTCTTCC
GGAAGGGGAGGAAGCGGGAGGCAGGCCTGCACCTGATGTGGAGGTACAGGCAGATCTTCCCTACCGGGGAGGGATGG
ATGGTTGGATACAGGTGGCTGGGCTATTCCATCCATCTGGAAGCACATTTGAGCCTCCAGGCTTCTCCTTGACGTCA
TTCCTCTCCTTCCTTGCTGCAAAATGGCTCTGCACCAGCCGGCCCCCCAGGAGGTCTGGCAGAGCTGAGAGCCATGGC
CTGCAGGGGCTCCATATGTCCCTACGCGTGCAGCAGGCAAACAAGA

FIGURE 33

```
CTCTGTCCTGGGACTTGGTGGTGCTACCCTTGGCCTCCCACAGTCCTGCCACCCTGCTGCCGCCACC**ATG**CTGCCCC
CTGGGACTGCGACCCTCTTGACTCTGCTCCTGGCAGCTGGCTCGCTGGGCCAGAAGCCTCAGAGGCCACGCCGGCCC
GCATCCCCCATCAGCACCATCCAGCCCAAGGCCAATTTTGATGCTCAGCAGTTTGCAGGGACCTGGCTCCTTGTGGC
TGTGGGCTCCGCTTGCCGTTTCCTGCAGGAGCAGGGCCACGGGCCGAGGCCACCACACTGCATGTGGCTCCCCAGG
GCACAGCCATGGCTGTCAGTACCTTCCGAAAGCTGGATGGGATCTGCTGGCAGGTGCGCCAGCTCTATGGAGACACA
GGGGTCCTCGGCCGCTTCCTGCTTCAAGCCCGAGGCGCCCGAGGGGCTGTGAACGTGGTTGTCGCTGAGACTGACTA
CCAGAGTTTCGCTGTCCTGTACCTGGAGCGGGCGGGGCAGCTGTCAGTGAAGCTCTACGCCCGCTCGCTCCCTGTGA
GCGACTCGGTCCTGAGTGGGTTTGAGCAGCGGGTCCAGGAGGCCCACCTGACTGAGGACCAGATCTTCTACTTCCCC
AAGTACGGCTTCTGCGAGGCTGCAGACCAGTTCCACGTCCTGGACGAAGTGAGGAGG**TGA**GGCCGGCACACAGCTCC
AGTGCTGAGAAGTCAGTGCCCCGAGAGACGACCCCACCAGTGGGGTGCCCGCTGCCTGTCCTCCGTGAAACCAGCCT
CAGATCAGGGCCCTGCCACCCAGGGCAGGGGATCTTCTGCCGGCTGCCCCAGAGGACAGTGGGTGGAGTGGTACCTA
CTTATTAAATGTCTCAGACCCCAAAAAAAAAAAAAAAAAAA
```

FIGURE 34

```
AAGAACAATTGTCTCTGGACGGCAGCTATGCGACTCACCGTGCTGTGTGCTGTGTGCCTGCTGCCTGGCAGCCTGGC
CCTGCCGCTGCCTCAGGAGGCGGGAGGCATGAGTGAGCTACAGTGGGAACAGGCTCAGGACTATCTCAAGAGATTTT
ATCTCTATGACTCAGAAACAAAAAATGCCAACAGTTTAGAAGCCAAACTCAAGGAGATGCAAAAATTCTTTGGCCTA
CCTATAACTGGAATGTTAAACTCCCGCGTCATAGAAATAATGCAGAAGCCCAGATGTGGAGTGCCAGATGTTGCAGA
ATACTCACTATTTCCAAATAGCCCAAAATGGACTTCCAAAGTGGTCACCTACAGGATCGTATCATATACTCGAGACT
TACCGCATATTACAGTGGATCGATTAGTGTCAAAGGCTTTAAACATGTGGGCAAAGAGATCCCCCTGCATTTCAGG
AAAGTTGTATGGGGAACTGCTGACATCATGATTGGCTTTGCGCGAGGAGCTCATGGGGACTCCTACCCATTTGATGG
GCCAGGAAACACGCTGGCTCATGCCTTTGCGCCTGGGACAGGTCTCGGAGGAGATGCTCACTTCGATGAGGATGAAC
GCTGGACGGATGGTAGCAGTCTAGGGATTAACTTCCTGTATGCTGCAACTCATGAACTTGGCCATTCTTTGGGTATG
GGACATTCCTCTGATCCTAATGCAGTGATGTATCCAACCTATGGAAATGGAGATCCCCAAAATTTTAAACTTTCCCA
GGATGATATTAAAGGCATTCAGAAACTATATGGAAAGAGAAGTAATTCAAGAAAGAAATAGAAACTTCAGGCAGAAC
ATCCATTCATTCATTCATTGGATTGTATATCATTGTTGCACAATCAGAATTGATAAGCACTGTTCCTCCACTCCATT
TAGCAATTATGTCACCCTTTTTTATTGCAGTTGGTTTTTGAATGTCTTTCACTCCTTTTATTGGTTAAACTCCTTTA
TGGTGTGACTGTGTCTTATTCCATCTATGAGCTTTGTCAGTGCGCGTAGATGTCAATAAATGTTACATACACAAATA
AATAAAATATTTAGGCCATGGTAAATTTACC
```

FIGURE 35

GAGCGAGCACCTTCGACGCGGTCCGGGGACCCCCTCGTCGCTGTCCTCCCGACGCGGACCCGCGTGCCCCAGGCCTC
GCGCTGCCCGGCCGGCTCCTCGTGTCCCACTCCCGGCGCACGCCCTCCCGCGCCCCTCTTCTCGGCGCGCGCGCAGC
ATGGCGCCCCCGCAGGTCCTCGCGTTCGGGCTTCTGCTTGCCGCGGCGACGGCGACTTTTGCCGCAGCTCAGGAAGA
ATGTGTCTGTGAAAACTACAAGCTGGCCGTAAACTGCTTTGTGAATAATAATCGTCAATGCCAGTGTACTTCAGTTG
GTGCACAAAATACTGTCATTTGCTCAAAGCTGGCTGCCAAATGTTTGGTGATGAAGGCAGAAATGAATGGCTCAAAA
CTTGGGAGAAGAGCAAAACCTGAAGGGGCCCTCCAGAACAATGATGGGCTTTATGATCCTGACTGCGATGAGAGCGG
GCTCTTTAAGGCCAAGCAGTGCAACGGCACCTCCACGTGCTGGTGTGTGAACACTGCTGGGGTCAGAAGAACAGACA
AGGACACTGAAATAACCTGCTCTGAGCGAGTGAGAACCTACTGGATCATCATTGAACTAAAACACAAAGCAAGAGAA
AAACCTTATGATAGTAAAAGTTTGCGGACTGCACTTCAGAAGGAGATCACAACGCGTTATCAACTGGATCCAAAATT
TATCACGAGTATTTTGTATGAGAATAATGTTATCACTATTGATCTCGGTTCAAAATTCTTCTCAAAAAACTCAGAATG
ATGTGGACATAGCTGATGTGGCTTATTATTTTGAAAAAGATGTTAAAGGTGAATCCTTGTTTCATTCTAAGAAAATG
GACCTGACAGTAAATGGGGAACAACTGGATCTGGATCCTGGTCAAACTTTAATTTATTATGTTGATGAAAAAGCACC
TGAATTCTCAATGCAGGGTCTAAAAGCTGGTGTTATTGCTGTTGTTATTGTGGTTGTGGTGATGGCAGTTGTTGCTGGAA
TTGTTGTGCTGGTTATTTCCAGAAAGAAGAGAATGGCAAAGTATGAGAAGGCTGAGATAAAGGAGATGGGTGAGATG
CATAGGGAACTCAATGCA**TAA**CTATATAATTTGAAGATTATAGAAGAAGGGAAATAGCAAATGGACACAAATTACAA
ATGTGTGTGCGTGGGACGAAGACATCTTTGAAGGTCATGAGTTTGTTAGTTTAACATCATATATTTGTAATAGTGAA
ACCTGTACTCAAAATATAAGCAGCTTGAAACTGGCTTTACCAATCTTGAAATTTGACCACAAGTGTCTTATATATGC
AGATCTAATGTAAAATCCAGAACTTGGACTCCATCGTTAAAATTATTTATGTGTAACATTCAAATGTGTGCATTAAA
TATGCTTCCACAGTAAAATCTGAAAAACTGATTTGTGATTGAAAGCTGCCTTTCTATTTACTTGAGTCTTGTACATA
CATACTTTTTTATGAGCTATGAAATAAAACATTTTAAACTG

FIGURE 36

GCTCCGGGCTGAAGATTGCTTCTCTTCTCTCCTCCAAGGTCTAGTGACGGAGCCCGCGCGCGCGCCACC**ATG**CGGCA
GAAGGCGGTATCCGTTTTCTTGTGCTACCTGCTGCTCTTCACTTGCAGTGGGGTGGAGGCAGGTAAGAAAAAGTGCT
CGGAGAGCTCGGACAGCGGCTCCGGGTTCTGGAAGGCCCTGACCTTCATGGCCGTCGGAGGAGGACTCGCAGTCGCC
GGGCTGCCCGCGCTGGGCTTCACCGGCGCCGGCATCGCGGCCAACTCGGTGGCTGCCTCGCTGATGAGCTGGTCTGC
GATCCTGAATGGGGGCGGCGTGCCCGCCGGGGGGGCTAGTGGCCACGCTGCAGAGCCTCGGGGCTGGTGGCAGCAGCG
TCGTCATAGGTAATATTGGTGCCCTGATGCGGTACGCCACCCACAAGTATCTCGATAGTGAGGAGGATGAGGAG**TAG**
CCAGCAGCTCCCAGAACCTCTTCTTCCTTCTTGGCCTAACTCTTCCAGTTAGGATCTAGAACTTTGCCTTTTTTTTT
TTTTTTTTTTTTTGAGATGGGTTCTCACTATATTGTCCAGGCTAGAGTGCAGTGGCTATTCACAGATGCGAACATA
GTACACTGCAGCCTCCAACTCCTAGCCTCAAGTGATCCTCCTGTCTCAACCTCCCAAGTAGGATTACAAGCATGCGC
CGACGATGCCCAGAATCCAGAACTTTGTCTATCACTCTCCCCAACAACCTAGATGTGAAAACAGAATAAACTTCACC
CAGAAAA

FIGURE 37

GCCGCCGAGGGCAGCCAGCCCCTCCCCTACCCGGAGCAGCCCGCTGGGGCCGTCCCGAGCGGCGACACACTAGGAGT
CCCGGCCGGCCAGCCAGGCAGCCGCGGTCCCGGGACTCGGCCGTGAGTGCTGCGGGACGGATGGTGGCGGCGGGAG
CGCGGAGACCACGGCGGGCGCCGTGGAGCCGGGCGCCGTGCAGCCGGAGCTGCGCGCGGGGCATGCGGCTGCGCCCC
GGCCCCTCGGCCCCCGGCCTCGGCCCCCGCGCTCCGGCCCCAGCCCCGGCCGCCGCCGCCCCCGCGGAGTGCAGCGACC
GCGCCGCCGCTGAGGGAGGCGCCCCAC**ATG**CGCGGGCCCCGGCGCCGCTGTACGCCTGCCTCCTGGGGCTCTGCG
CGCTCCTGCCCCGGCTCGCAGGTCTCAACATATGCACTAGTGGAAGTGCCACCTCATGTGAAGAATGTCTGCTAATC
CACCCAAAATGTGCCTGGTGCTCCAAAGAGGACTTCGGAAGCCCACGGTCCATCACCTCTCGGTGTGATCTGAGGGC
AAACCTTGTCAAAAATGGCTGTGGAGGTGAGATAGAGAGCCCAGCCAGCAGCTTCCATGTCCTGAGGAGCCTGCCCC
TCAGCAGCAAGGGTTCGGGCTCTGCAGGCTGGGACGTCATTCAGATGACACCACAGGAGATTGCCGTGAACCTCCGG
CCCGGTGACAAGACCACCTTCCAGCTACAGGTTCGCCAGGTGGAGGACTATCCTGTGGACCTGTACTACCTGATGGA
CCTCTCCCTGTCCATGAAGGATGACTTGGACAATATCCGGAGCCTGGGCACCAAACTCGCGGAGGAGATGAGGAAGC
TCACCAGCAACTTCCGGTTGGGATTTGGGTCTTTTGTTGATAAGGACATCTCTCCTTTCTCCTACACGGCACCGAGG
TACCAGACCAATCCGTGCATTGGTTACAAGTTGTTTCCAAATTGCGTCCCCTCCTTTGGGTTCCGCCATCTGCTGCC
TCTCACAGACAGAGTGGACAGCTTCAATGAGGAAGTTCGGAAACAGAGGGTGTCCCGGAACCGAGATGCCCCTGAGG
GGGGCTTTGATGCAGTACTCCAGGCAGCCGTCTGCAAGGAGAAGATTGGCTGGCGAAAGGATGCACTGCATTGCTG
GTGTTCACAACAGATGATGTGCCCCACATCGCATTGGATGGAAAATTGGGAGGCCTGGTGCAGCCACACGATGGCCA
GTGCCACCTGAACGAGGCCAACGAGTACACAGCATCCAACCAGATGGACTATCCATCCCTTGCCTTGCTTGGAGAGA
AATTGGCAGAGAACAACATCAACCTCATCTTTGCAGTGACAAAAACCATTATATGTGTACAAGAATTTTACAGCC
CTGATACCTGGAACAACGGTGGAGATTTAGATGGAGACTCCAAAAATATTATTCAACTGATTATTAATGCATACAA
TAGTATCCGGTCTAAAGTGGAGTTGTCAGTCTGGGATCAGCCTGAGGATCTTAATCTCTTCTTTACTGCTACCTGCC
AAGATGGGGTATCCTATCCTGGTCAGAGGAAGTGTGAGGGTCTGAAGATTGGGGACACGGCATCTTTGAAGTATCA
TTGGAGGCCCGAAGCTGTCCCAGCAGACACACGGAGCATGTGTTTGCCCTGCGGCCGGTGGGATTCCGGGACAGCCT
GGAGGTGGGGGTCACCTACAACTGCACGTGCGGCTGCAGCGTGGGGCTGGAACCCAACAGCGCCAGGTGCAACGGGA
GCGGGACCTATGTCTGCGGCCTGTGTGAGTGCAGCCCCGGCTACCTGGGCACCAGGTGCGAGTGCCAGGATGGGGAG
AACCAGAGCGTGTACCAGAACCTGTGCCGGGAGGCAGAGGGCAAGCCACTGTGCACGGGCGTGGGGACTGCAGCTG
CAACCAGTGCTCCTGCTTCGAGAGCGAGTTTGGCAAGATCTATGGGCCTTTCTGTGAGTGCGACAACTTCTCCTGTG
CCAGGAACAAGGGAGTCCTCTGCTCAGGCCATGGCGAGTGTCACTGCGGGGAATGCAAGTGCCATGCAGGTTACATC
GGGGACAACTGTAACTGCTCGACAGACATCAGCACATGCCGGGGCAGAGATGGCCAGATCTGCAGCGAGCGTGGGCA
CTGTCTCTGTGGGCAGTGCCAATGCACGGAGCCGGGGGCCTTTGGGGAGATGTGTGAGAAGTGCCCCCACCTGCCCGG
ATGCATGCAGCACCAAGAGAGATTGCGTCGAGTGCCTGCTGCTCCACTCTGGGAAACCTGACAACCAGACCTGCCAC
AGCCTATGCAGGGATGAGGTGATCACATGGGTGGACACCATCGTGAAAGATGACCAGGAGGCTGTGCTATGTTTCTA
CAAAACCGCCAAGGACTGCGTCATGATGTTCACCTATGTGGAGCTCCCCAGTGGGAAGTCCAACCTGACCGTCCTCA
GGGAGCCAGAGTGTGGAAACACCCCCAACGCCATGACCATCCTCCTGGCTGTGGTCGGTAGCATCCTCCTTGTTGGG
CTTGCACTCCTGGCTATCTGGAAGCTGCTTGTCACCATCCACGACCGGAGGGAGTTTGCAAAGTTTCAGAGCGAGCG
ATCCAGGGCCCGCTATGAAATGGCTTCAAATCCATTATACAGAAAGCCTATCTCCACGCACACTGTGGACTTCACCT
TCAACAAGTTCAACAAATCCTACAATGGCACTGTGGAC**TGA**TGTTTCCTTCTCCGAGGGGCTGGAGCGGGGATCTGA
TGAAAAGGTCAGACTGAAACGCCTTGCACGGCTGCTCGGCTTGATCACAGCTCCCTAGGTAGGACCACCACAGAGAAGA
CCTTCTAGTGAGCCTGGGCCAGGAGCCCACAGTGCTGTACAACAAGGGAAAGGTAGCCTGGCCATGTCACCTGGCTG
CTAGCCAGAGCCATGCCAGGTTCGCGTCCCTAAGAGCTTGGGATAAAGCAAGGGGACCTTGGCGCTCTCAGCTTTCC
CTGCCACATCCAGCTTGTTGTCCCAATGAAATACTGAGATGCTGGGCTGTCTCTCCCTTCCAGGAATCGTGGGCCCC
CAGCCTGGCCAGACAAGAAGACTGTCAGGAAGGGTCGGAGTCTGTAAAACCAGCATACAGTTTGGCTTTTTTCACAT
TGATCATTTTTATATGAAATAAAAAGATCCTGCATTTATGGTGTAGTTCTGAGTCCTGAGACTTTTCTGCGTGATGC
TATGCCTTGCACACAGGTGTTGGTGATGGGGCTGTTGAGATGCCTGTTGAAGGTACATCGTTTGCAAATGTCAGTTT
CCTCTCCTGTCCGTGTTTGTTTAGTACTTTTATAATGAAAAGAAACAAGATTGTTTGGGATTGGAAGTAAAGATTAA
AACCAAAAGAATTTGTGTTTGTCTGCCC

FIGURE 38

CGACTCACTATAGGGCGAATTGAATTTAGCGGCCGCGAATTCGCCCT**ATG**CTGCCACAAATACCCTTTTTGCTGCT
AGTATCCTTGAACTTGGTTCATGGAGTGTTTTACGCTGAACGATACCAAATGCCCACAGGCATAAAAGGCCCACTAC
CCAACACCAAGACACAGTTCTTCATTCCCTACACCATAAAGAGTAAAGGTATAGCAGTAAGAGGAGAGCAAGGTACT
CCTGGTCCACCAGGCCCTGCTGGACCTCGAGGGCACCCAGGTCCTTCTGGACCACCAGGAAAACCAGGCTACGGAAG
TCCTGGACTCCAAGGAGAGCCAGGGTTGCCAGGACCACCGGGACCATCAGCTGTAGGGAAACCAGGTGTGCCAGGAC
TCCCAGGAAAACCAGGAGAGAGGACCCATATGGACCAAAAGGAGATGTTGGACCAGCTGGCCTACCAGGACCCCGG
GGCCCACCAGGACCACCTGGAATCCCTGGACCGGCTGGAATTTCTGTGCCAGGAAAACCTGGACAACAGGGACCCAC
AGGAGCCCCAGGACCCAGGGGCTTTCCTGGAGAAAAGGGTGCACCAGGAGTCCCTGGTATGAATGGACAGAAAGGGG
AAATGGGATATGGTGCTCCTGGTCGTCCAGGTGAGAGGGGTCTTCCAGGCCCTCAGGGTCCCACAGGACCATCTGGC
CCTCCTGGAGTGGGAAAAAGAGGTGAAAATGGGGTTCCAGGACAGCCCAGGCATCAAAGGTGATAGAGGTTTTCCGGG
AGAAATGGGACCAATTGGCCCACCAGGTCCCCAAGGCCCTCCTGGGGAACGAGGGCCAGAAGGCATTGGAAAGCCAG
GAGCTGCTGGAGCCCCAGGCCAGCCAGGGATTCCAGGAACAAAAGGTCTCCCTGGGGCTCCAGGAATAGCTGGGCCC
CCAGGGCCTCCTGGCTTTGGGAAACCAGGCTTGCCAGGCCTGAAGGGAGAAAGAGGACCTGCTGGCCTTCCTGGGGG
TCCAGGTGCCAAAGGGGAACAAGGGCCAGCAGGTCTTCCTGGGAAGCCAGGTCTGACTGGACCCCCTGGGAATATGG
GACCCCAAGGACCAAAAGGCATCCCGGGTAGCCATGGTCTCCCAGGCCCTAAAGGTGAGACAGGGCCAGCTGGGCCT
GCAGGATACCCTGGGGCTAAGGGTGAAAGGGGTTCCCCTGGGTCAGATGGAAAACCAGGGTACCCAGGAAAACCAGG
TCTCGATGGTCCTAAGGGTAACCCAGGGTTACCAGGTCCAAAAGGTGATCCTGGAGTTGGAGGACCTCCTGGTCTCC
CAGGCCCTGTGGGCCCAGCAGGAGCAAAGGGAATGCCCGGACACAATGGAGAGGCTGGCCCAAGAGGTGCCCCTGGA
ATACCAGGTACTAGAGGCCCTATTGGGCCACCAGGCATTCCAGGATTCCCTGGGTCTAAAGGGGATCCAGGAAGTCC
CGGTCCTCCTGGCCCAGCTGGCATAGCAACTAAGGGCCTCAATGGACCCACCGGGCCACCAGGGCCTCCAGGTCCAA
GAGGCCACTCTGGAGAGCCTGGTCTTCCAGGGCCCCCTGGGCCTCCAGGCCCACCAGGTCAAGCAGTCATGCCTGAG
GGTTTTATAAAGGCAGGCCAAAGGCCCAGTCTTTCTGGGACCCCTCTTGTTAGTGCCAACCAGGGGGGTAACAGGAAT
GCCTGTGTCTGCTTTTACTGTTATTCTCTCCAAAGCTTACCCAGCAATAGGAACTCCCATACCATTTGATAAAATTT
TGTATAACAGGCAACAGCATTATGACCCAAGGACTGGAATCTTTACTTGTCAGATACCAGGAATATACTATTTTTCA
TACCACGTGCATGTGAAAGGGACTCATGTTTGGGTAGGCCTGTATAAGAATGGCACCCCTGTAATGTACACCTATGA
TGAATACACCAAAGGCTACCTGGATCAGGCTTCAGGGAGTGCCATCATCGATCTCACAGAAAATGACCAGGTGTGGC
TCCAGCTTCCCAATGCCGAGTCAAATGGCCTATACTCCTCTGAGTATGTCCACTCCTCTTTCTCAGGATTCCTAGTG
GCTCCAATG**TGA**GTACAAGGGCGAATTCGTTTAAACCTGCAGGACTAGTCCCTTTAGTGAGGGTTAATT

FIGURE 39

GTTGGTGACCAAGAGTACATCTCTTTTCAAATAGCTGGATTAGGTCCTC**ATG**CTGCTGTGGTCATTGCTGGTCATCT
TTGATGCAGTCACTGAACAGGCAGATTCGCTGACCCTTGTGGCGCCCTCTTCTGTCTTCGAAGGAGACAGCATCGTT
CTGAAATGCCAGGGAGAACAGAACTGGAAAATTCAGAAGATGGCTTACCATAAGGATAACAAAGAGTTATCTGTTTT
CAAAAAATTCTCAGATTTCCTTATCCAAAGTGCAGTTTTAAGTGACAGTGGTAACTATTTCTGTAGTACCAAAGGAC
AACTCTTTCTCTGGGATAAAACTTCAAATATAGTAAAGATAAAAGTCCAAGAGCTCTTTCAACGTCCTGTGCTGACT
GCCAGCTCCTTCCAGCCCATCGAAGGGGGTCCAGTGAGCCTGAAATGTGAGACCCGGCTCTCTCCACAGAGGTTGGA
TGTTCAACTCCAGTTCTGCTTCTTCAGAGAAAACCAGGTCCTGGGGTCAGGCTGGAGCAGCTCTCCGGAGCTCCAGA
TTTCTGCCGTGTGGAGTGAAGACACAGGGTCTTACTGGTGCAAGGCAGAAACGGTGACTCACAGGATCAGAAAACAG
AGCCTCCAATCCCAGATTCACGTGCAGAGAATCCCCATCTCTAATGTAAGCTTGGAGATCCGGGCCCCCGGGGGACA
GGTGACTGAAGGACAAAAACTGATCCTGCTCTGCTCAGTGGCTGGGGGTACAGGAAATGTCACATTCTCCTGGTACA
GAGAGGCCACAGGAACCAGTATGGGAAAGAAAACCCAGCGTTCCCTGTCAGCAGAGCTGGAGATCCCAGCTGTGAAA
GAGAGTGATGCCGGCAAATATTACTGTAGAGCTGACAACGGCCATGTGCCTATCCAGAGCAAGGTGGTGAATATCCC
TGTGAGAATTCCAGTGTCTCGCCCTGTCCTCACCCTCAGGTCTCCTGGGGCCCAGGCTGCAGTGGGGGACCTGCTGG
AGCTTCACTGTGAGGCCCTGAGAGGCTCTCCCCCAATCTTGTACCAATTTTATCATGAGGATGTCACCCTTGGGAAC
AGCTCGGCCCCCTCTGGAGGAGGGGGCTCCTTCAACCTCTCTTTGACTGCAGAACATTCTGGAAACTACTCCTGTGA
GGCCAACAACGGCCTGGGGGCCCAGTGCAGTGAGGCAGTGCCAGTCTCCATCTCAGGACCTGATGGCTATAGAAGAG
ACCTCATGACAGCTGGAGTTCTCTGGGGACTGTTGGTGTCCTTGGTTTCACTGGTGTTGCTTTGCTGTTGTATGCC
TTGTTCCACAAGATATCAGGAGAAAGTTCTGCCACTAATGAACCCAGAGGGGCTTCCAGGCCAAATCCTCAAGAGTT
CACCTATTCAAGCCCAACCCCAGACATGGAGGAGCTGCAGCCAGTGTATGTCAATGTGGGCTCTGTAGATGTGGATG
TGGTTTATTCTCAGGTCTGGAGCATGCAGCAGCCAGAAAGCTCAGCAAACATCAGGACACTTCTGGAGAACAAGGAC
TCCCAAGTCATCTACTCTTCTGTGAAGAAATCA**TAA**CACTTGGAGGAATCAGAAGGGAAGATCAACAGCAAGGATGG
GGCATCATTAAGACTTGCTATAAAACCTTATGAAAATGCTTGAGGCTTATCACCTGCCACAGCCAGAACGTGCCTCA
GGAGGCACCTCCTGTCATTTTTGTCCTGATGATGTTTCTTCTCCAATATCTTCTTTTACCTATCAATATTCATTGAA
CTGCTGCTACATCCAGACACTGTGCAAATAAATTATTTCTGCTACCTTCAAAAAAAAAAAAAAAAAAAAAAATGCAG

FIGURE 40

GGCACGAGG**CTG**CGCCAGGGCCTGAGCGGAGGCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTC
ACTGAGCCAGCGCCTGCGCCTCTACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTCACCTGCTTCCTCC
TGGGCGTGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCGCACTGTCCTCTGCATCGACTTCATGGTTTTC
ACGGTGCCGGCTGCTTCACATCTTCACGGTCAACAAACAGCTGGGGCCCAAGATCGTCATCGTGAGCAAGATGATGAA
GGACGTGTTCTTCTTCCTCTTCTTCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGCTCCTGAGGC
CACGGGACAGTGACTTCCCAAGTATCCTGCGCCGCGCTCTTCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTCCC
CAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCTTCTGGGCACACCCTCCTGG
GGCCCAGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGCTGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGTGG
CCAACATCCTGCTGGTCAACTTGCTCATTGCCATGTTCAGTTACACATTCGGCAAAGTACAGGGCAACAGCGATCTC
TACTGGAAGGCGCAGCGTTACCGCCTCATCCGGGAATTCCACTCTCGGCCCGCGCTGGCCCCGCCCTTTATCGTCAT
CTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGACCCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCGAGC
ATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGCGGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTG
CTGGCACGCGCTAGGGACAAGCGGGAGAGCGACTCCGAGCGTCTGAAGCGCACGTCCCAGAAGGTGGACTTGGCACT
GAAACAGCTGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGGTCCAGCAGTGTAGCCGCG
TCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGCTGCCCCCCAGGTGGGCCGCCACCCCTGACCTGCCT
GGGTCCAAAGAC**TGA**GCCCTGCTGGCGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGCTC
ATCTGGGCCTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCACTGTCAGGAC
CACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGGCTCCTCCCAGAACCAGTCCCAGCCTGGGAGGATCA
AGGCCTGGATCCCGGGCCGTTATCCATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCA
CTCACAGATTCCTCACACTGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAAA

FIGURE 41

GAGAGAACAGCGTGAGCCTGTGTGCTTGTGTGCTGAGCCCTCATCCCCTCCTGGGGCCAGGCTTGGGTTTCACCTGC
AGAATCGCTTGTGCTGGGCTGCCTGGGCTGTCCTCAGTGGCACCTGCATGAAGCCGTTCTGGCTGCCAGAGCTGGAC
AGCCCCAGGAAAACCCACCTCTCTGCAGAGCTTGCCCAGCTGTCCCCGGGAAGCCAAATGCCTCTCATGTAAGTCTT
CTGCTCGACGGGGTGTCTCCTAAACCCTCACTCTTCAGCCTCTGTTTGACCATGAAATGAAGTGACTGAGCTCTATT
CTGTACCTGCCACTCTATTTCTGGGGTGACTTTTGTCAGCTGCCCAGAATCTCCAAGCCAGGCTGGTTCTCTGCATC
CTTTCAATGACCTGTTTTCTTCTGTAACCACAGGTTCGGTGGTGAGAGGAAGCCTCGCAGAATCCAGCAGAATCCTC
ACAGAATCCAGCAGCAGCTCTGCTGGGGACATGGTCATGGTGCAACCCACAGCAAAGCCCTGACCTGACCTCCTGA
TGCTCAGGAGAAGCC**ATG**GGCCCCTCCTGTCCTGTGTTCCTGTCCTTCACAAAGCTCAGCCTGTGGTGGCTCCTTCT
GACCCCAGCAGGTGGAGAGGAAGCTAAGCGCCCACCTCCCAGGGGCTCCTGGAGACCCACTCTCCTCTCCCAGTCCCA
CGGCATTGCCGCAGGGAGGCTCGCATACCGAGACTGAGGACCGGCTCTTCAAACACCTCTTCCGGGGCTACAACCGC
TGGGCGCGCCCGGTGCCCAACACTTCAGACGTGGTGATTGTGCGCTTTGGACTGTCCATCGCTCAGCTCATCGATGT
GGATGAGAAGAACCAAATGATGACCACCAACGTCTGGCTAAAACAGGAGTGGAGCGACTACAAACTGCGCTGGAACC
CCGCTGATTTTGGCAACATCACATCTCTCAGGGTCCCTTCTGAGATGATCTGGATCCCCGACATTGTTCTCTACAAC
AATGCAGATGGGGAGTTTGCAGTGACCCACATGACCAAGGCCCACCTCTTCTCCACGGGCACTGTGCACTGGGTGCC
CCCGGCCATCTACAAGAGCTCCTGCAGCATCGACGTCACCTTCTTCCCCCTTCGACCAGCAGAACTGCAAGATGAAGT
TTGGCTCCTGGACTTATGACAAGGCCAAGATCGACCTGGAGCAGATGGAGCAGACTGTGGACCTGAAGGACTACTGG
GAGAGCGGCGAGTGGGCCATCGTCAATGCCACGGGCACCTACAACAGCAAGAAGTACGACTGCTGCGCCGAGATCTA
CCCCGACGTCACCTACGCCTTCGTCATCCGGCGGCTGCCGCTCTTCTACACCATCAACCTCATCATCCCCTGCCTGC
TCATCTCCTGCCTCACTGTGCTGGTCTTCTACCTGCCCTCCGACTGCGGCGAGAAGATCACGCTGTGCATTCCGGTG
CTGCTGTCACTCACCGTCTTCCTGCTGCTCATCACTGAGATCATCCCGTCCACCTCGCTGGTCATCCCGCTCATCGG
CGAGTACCTGCTGTTCACCATGATCTTCGTCACCCTGTCCATCGTCATCACCGTCTTCGTGCTCAATGTGCACCACC
GCTCCCCCAGCACCCACACCATGCCCCACTGGGTGCGGGGGGCCCTTCTGGGCTGTGTGCCCCGGTGGCTTCTGATG
AACCGGCCCCCACCACCCGTGGAGCTCTGCCACCCCCTACGCCTGAAGCTCAGCCCCTCTTATCACTGGCTGGAGAG
CAACGTGGATGCCGAGGAGAGGGAGGTGGTGGTGGAGGAGGAGGACAGATGGGCATGTGCAGGTCATGTGGCCCCCT
CTGTGGGCACCCTCTGCAGCCACGGCCCACCTGCACTCTGGGGCCTCAGGTCCCAAGGCTGAGGCTCTGCTGCAGGAG
GGTGAGCTGCTGCTATCACCCCACATGCAGAAGGCACTGGAAGGTGTGCACTACATTGCCGACCACCTGCGGTCTGA
GGATGCTGACTCTTCGGTGAAGGAGGACTGGAAGTATGTTGCCATGGTCATCGACAGGATCTTCCTCTGGCTGTTTA
TCATCGTCTGCTTCCTGGGGACCATCGGCCTCTTTCTGCCTCCGTTCCTAGCTGGAATGATC**TGA**CTGCACCTCCCT
CGAGCTGGCTCCCAGGGCAAAGGGGAGGGTTCTTGGATGTGGAAGGGCTTTGAACAATGTTTAGATTTGGAGATGAG
CCCAAAGTGCCAGGGAGGAACAGCCAGGTGAGGTGGGAGGTTGGAGAGCCAGGTGAGGTCTCTCTAAGTCAGGCTGGG
GTTGAAGTTTGGAGTCTGTCCGAGTTTGCAGGGTGCTGAGCTGTATGGTCCAGCAGGGGAGTAATAAGGGCTCTTCC
GGAAGGGGAGGAAGCGGGAGGCAGGCCTGCACCTGATGTGGAGGTACAGGCAGATCTTCCCTACCGGGGAGGGATGG
ATGGTTGGATACAGGTGGCTGGGCTATTCCATCCATCTGGAAGCACATTTGAGCCTCCAGGCTTCTCCTTGACGTCA
TTCCTCTCCTTCCTTGCTGCAAAATGGCTCTGCACCAGCCGGCCCCCAGGAGGTCTGGCAGAGCTGAGAGCCATGGC
CTGCAGGGGCTCCATATGTCCCTACGCGTGCAGCAGGCAAACAAGA

FIGURE 42

CCGAGACTCACGGTCAAGCTAAGGCGAAGAGTGGGTGGCTGAAGCCATACTATTTTATAGAATTA**ATG**GAAAGCAGA
AAAGACATCACAAACCAAGAAGAACTTTGGAAAATGAAGCCTAGGAGAAATTTAGAAGAAGACGATTATTTGCATAA
GGACACGGGAGAGACCAGCATGCTAAAAAGACCTGTGCTTTTGCATTTGCACCAAACAGCCCATGCTGATGAATTTG
ACTGCCCTTCAGAACTTCAGCACACACAGGAACTCTTTCCACAGTGGCACTTGCCAATTAAAATAGCTGCTATTATA
GCATCTCTGACTTTTCTTTACACTCTTCTGAGGGAAGTAATTCACCCTTTAGCAACTTCCCATCAACAATATTTTTA
TAAAATTCCAATCCTGGTCATCAACAAAGTCTTGCCAATGGTTTCCATCACTCTCTTGGCATTGGTTTACCTGCCAG
GTGTGATAGCAGCAATTGTCCAACTTCATAATGGAACCAAGTATAAGAAGTTTCCACATTGGTTGGATAAGTGGATG
TTAACAAGAAAGCAGTTTGGGCTTCTCAGTTTCTTTTTTGCTGTACTGCATGCAATTTATAGTCTGTCTTACCCAAT
GAGGCGATCCTACAGATACAAGTTGCTAAACTGGGCATATCAACAGGTCCAACAAAATAAAGAAGATGCCTGGATTG
AGCATGATGTTTGGAGAATGGAGATTTATGTGTCTCTGGGAATTGTGGGATTGGCAATACTGGCTCTGTTGGCTGTG
ACATCTATTCCATCTGTGAGTGACTCTTTGACATGGAGAGAATTTCACTATATTCAGAGCAAGCTAGGAATTGTTTC
CCTTCTACTGGGCACAATACACGCATTGATTTTTGCCTGGAATAAGTGGATAGATATAAAACAATTTGTATGGTATA
CACCTCCAACTTTTATGATAGCTGTTTTCCTTCCAATTGTTGTCCTGATATTTAAAAGCATACTATTCCTGCCATGC
TTGAGGAAGAAGATACTGAAGATTAGACATGGTTGGGAAGACGTCACCAAAATTAACAAAACTGAGATATGTTCCCA
GTTG**TAG**AATTACTGTTTACACACATTTTTGTTCAATATTGATATATTTTATCACCAACATTTCAAGTTTGTATTTG
TTAATAAAATGATTATTCAAGGAAAAAAAAAAAAAAAAAAAA

FIGURE 43

```
GTGGCGCCCCCGGACGCCCACCCCGCAGTGGCCATGGAGAAGCCTTCGGGGCAGCCGTGGTGACCGTGTGGGACAGC
GATGCACAACACCACGGAGAAGCCCACCGATGCCTACGGAGAGCTGGACTTCACGGGGGCCGGCCGCAAGCACAGCA
ATTTCCTCCGGCTCTCTGACCGAACGGATCCAGCTGCAGTTTATAGTCTGGTCACACGCACATGGGGCTTCCGTGCC
CCGAACCTGGTGGTGTCAGTGCTGGGGGGATCGGGGGGCCCCGTCCTCCAGACCTGGCTGCAGGACCTGCTGCGTCG
TGGGCTGGTGCGGGCTGCCCAGAGCACAGGAGCCTGGATTGTCACTGGGGGTCTGCACACGGCATCGGCCGGCATG
TTGGTGTGGCTGTACGGGACCATCAGATGGCCAGCACTGGGGGCACCAAGGTGGTGGCCATGGGTGTGGCCCCCTGG
GGTGTGGTCCGGAATAGAGACACCCTCATCAACCCCAAGGGCTCGTTCCCTGCGAGGTACCGGTGGCGCGGTGACCC
GGAGGACGGGGTCCAGTTTCCCCTGGACTACAACTACTCGGCCTTCTTCCTGGTGGACGACGGCACACAGGCTGCC
TGGGGGGCGAGAACCGCTTCCGCTTGCGCCTGGAGTCCTACATCTCACAGCAGAAGACGGGCGTGGGAGGGACTGGA
ATTGACATCCCCTGTCCTGCTCCTCCTGATTGATGGTGATGAGAAGATGTTGACGCGAATAGAGAACGCCACCCAGGC
TCAGCTCCCATGTCTCCTCGTGGCTGGCTCAGGGGGAGCTGCGGACTGCCTGGCGGAGACCCTGGAAGACACTCTGG
CCCCAGGGAGTGGGGGAGCCAGGCAAGGCGAAGCCCGAGATCGAATCAGGCGTTTCTTTCCCAAAGGGGACCTTGAG
GTCCTGCAGGCCCAGGTGGAGAGGATTATGACCCGGAAGGAGCTCCTGACAGTCTATTCTTCTGAGGATGGGTCTGA
GGAATTCGAGACCATAGTTTTGAAGGCCCTTGTGAAGGCCTGTGGGAGCTCGGAGGCCTCAGCCTACCTGGATGAGC
TGCGTTTGGCTGTGGCTTGGAACCGCGTGGACATTGCCCAGAGTGAACTCTTTCGGGGGGGACATCCAATGGCGGTCC
TTCCATCTCGAAGCTTCCCTCATGGACGCCCTGCTGAATGACCGGCCTGAGTTCGTGCGCTTGCTCATTTCCCACGG
CCTCAGCCTGGGCCACTTCCTGACCCCGATGCGCCTGGCCCAACTCTACAGCGCGGCGCCCTCCAACTCGCTCATCC
GCAAACCTTTTGGACCAGGCGTCCCACAGCGCAGGCACCAAAGCCCCAGCCCTAAAAGGGGGGAGCTGCGGAGCTCCGG
CCCCCTGACGTGGGGCATGTGCTGAGGATGCTGCTGGGGAAGATGTGCGCGCCGAGGTACCCCTCCGGGGGCGCCTG
GGACCCTCACCCAGGCCAGGGCTTCGGGGAGAGCATGTATCTGCTCTCGGACAAGGCCACCTCGCCGCTCTCGCTGG
ATGCTGGCCTCGGGCAGGCCCCCTGGAGCGACCTGCTTCTTTGGGCACTGTTGCTGAACAGGGCACAGATGGCCATG
TACTTCTGGGAGATGGGTTCCAATGCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGGCACGCT
GGAGCCTGACGCTGAGGAGGCAGCACGGAGGAAAGACCTGGCGTTCAAGTTTGAGGGGATGGGCGTTGACCTCTTTG
GCGAGTGCTATCGCAGCAGTGAGGTGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCCCGCTCTGGGGGGATGCCACT
TGCCTCCAGCTGGCCATGCAAGCTGACGCCCGTGCCTTCTTTGCCCAGGATGGGGTACAGTCTCTGCTGACACAGAA
GTGGTGGGGAGATATGGCCAGCACTACACCCATCTGGGCCCTGGTTCTCGCCTTCTTTTGCCCTCCACTCATCTACA
CCCGCCTCATCACCTTCAGGAAATCAGAAGAGGAGCCCACACGGAGGAGCTAGAGTTTGACATGGATAGTGTCATT
AATGGGGAAGGGCCTGTCGGGACGGCGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTCC
GGGTTGCTGCGGGGGCCGCTGCGGGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGCGCCGGTGACCA
TCTTCATGGGCAACGTGGTCAGCTACCTGCTGTTCTTGCTGCTTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGCCG
GCGCCGCCCGGCTCCCTGGAGCTGCTGCTATTTCTGGGACCTTTCACGCTGCTGTGCGAGGAACTGCGCCAGGGCCT
GAGCGGAGGCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCT
ACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTGCACCTGCTTCCTCCTCGGGCGTGGGGCTGCCGGCTGACC
CCGGGTTTGTACCACCTGGGCCGCACTGTCCTCTGCATCGACTTCATGGTTTTCACGGTGCGGCTGCTTCACATCTT
CACGGTCAACAAACAGCTGGGGCCCAAGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTTCT
TCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGGCTCCTGAGGCCACGGGACAGTGACTTCCCAAGT
ATCCTGCGCCGCGTCTTCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTCCCCAGGAGGACATGGACGTGGCCCT
CATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCTTCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCT
CCCAGTATGCCAACTGGCTGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGGTGGCCAACATCCTGCTGGTCAACTTG
CTCATTGCCATGTTCAGTTACACAATTCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGAAGGCGCAGCGTTACCG
CCTCATCCGGGAATTCCACTCTCGGCCCCGCGCTGGCCCCGCCCTTTATCGTCATCTCCCACTTGCGCCTCCTGCTCA
GGCAATTGTGCAGGCGACCCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAG
GAAGCCGAGCGGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTGCTGGCACGCGGTAGGGACAAGCG
GGAGAGCGACTCCGAGCGTCTGGAGCGCACGTCCCAGAAGGTGGACTTGGCACTGAAACAGCTGGGACACATCCGCG
AGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGGGGTGGGTGACGTAGCC
GTTAGCAGCTCTGCCATGTTGCCCTCAGGTGGGCCGCCACCCCTTGACCTGCATGGGTCCAAAGAGTGAGCCATGCT
GGCGGATTTTAAGGAGAAGCCCCCACAGGGGATTTTGCTCTTAGAGTAAGGCTCATGTGGGCCTCGGCCCCCGCACC
TGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCACTGTCAGGACCACCTTTGGGAGTGTCATCCTTA
CAAACCACAGCATGCCCGGCTCCTCCCAGAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTAT
CCATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGATTCCTCACACTGGGG
AAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 44A

CGATAATTGAAAACCCAGATGTCCCACAGGATTTCGGGAATCAAGGGTCAACAGTAGAGTCCCTCTGTGATGATGTT
GTCTCTGTACTCACCGTCCTGTGTGAGAAGCTGCAAGCCGCCATAAATGACAGCCAGCAGCTGCAGCTTCTCTACCT
GGAGTGCATCCTGTCTGTGCTCAGCAGCTCCTCCTCCTCC**ATG**CACCTGCACAGGCGCTTCACGGACCTGATCTGGA
AAAACCTCTGCCCTGCTCTCATCGTGATCTTGGGGAATCCAATTCATGACAAAACCATCACCTCTGCTCACACCAGC
AGCACCAGTACCAGCCTGGAGTCGGACTCTGCGTCTCCGGGAGTGTCTGACCACGGCCGAGGATCAGGCTGCTCCTG
CACTGGGCCGGCCCTGAGCGGACCTGTGGCTCGGACTATCTATTACATCGCAGCCGAGCTGGTCCGGCTGGTGGGGT
CTGTGGACTCCATGAAGCCCGTGCTCCAGTCCCTCTACCACCGAGTGCTGCTCTACCCCCCACCCCAGCACCGGGTG
GAAGCCATCAAAATAATGAAAGAGATACTTGGGAGCCCACAGCGTCTCTGTGACTTGGCAGGACCCAGCTCCACTGA
ATCAGAGTCCAGAAAAAGATCAATTTCAAAAAGAAAGTCTCATCTGGATCTCCTCAAACTCATCATGGATGGCATGA
CCGAAGCATGCATCAAGGGTGGCATCGAAGCTTGCTATGCAGCCGTGTCCTGTGTCTGCACCTTGCTGGGTGCCCTG
GATGAGCTCAGCCAGGGGAAGGGCTTGAGCGAAGGTCAGGTGCAACTGCTGCTTCTGCGCCTTGAGGAGCTGAAGGA
TGGGGCTGAGTGGAGCCCGAGATTCCATGGAGATCAATGAGGCTGACTTCCGCTGGCAGCGGCGAGTGCTGTCCTCAG
AACACACGCCGTGGGAGTCAGGGAACGAGAGGGAGCCTTGACATCAGCATCAGTGTCACCACAGACACAGGCCAGACC
ACTCTCGAGGGAGAGTTGGGTCAGACTACACCCGAGGACCATTCGGGAAACCACAAGAACAGTCTCAAGTCGCCAGC
CATCCCAGAGGGTAAGGAGACGCTGAGCAAAGTATTGGAAACAGAGGCGGTAGACCAGCCAGATGTCGTGCAGAGAA
GCCACACGGTCCCTTACCCTGACATAACTAACTTCCTGTCAGTAGACTGCAGGACAAGGTCCTATGGATCTAGGTAT
AGTGAGAGCAATTTTAGCGTTGATGACCAAGACCTTTCTAGGACAGAGTTTGATTCCTGTGATCAGTACTCTATGGC
AGCAGAAAAGGACTCGGGCAGGTCCGACGTGTCAGACATTGGGTCGGACAACTGTTCACTAGCCGATGAAGAGCAGA
CACCCCGGGACTGCCTAGGCCACCGGTCCCTGCGAACTGCCGCCCTGTCTCTAAAACTGCTGAAGAACCAGGAGGCG
GATCAGCACAGCGCCAGGCTGTTCATACAGTCCCTGGAAGGCCTCCTCCCTCGGCTCCTGTCTCTCTCCAATGTAGA
GGAGGTGGACACCGCTCTGCAGAACTTTGCCTCTACTTTCTGCTCAGGCATGATGCACTCTCCTGGCTTTGACGGGA
ATAGCAGCCTCAGCTTCCAGATGCTGATGAACGCAGACAGCCTCTACACAGCTGCACACTGCGCCCTGCTCCTCAAC
CTGAAGCTCTCCCACGGTGACTACTACAGGAAGCGGCCGACCCTGGCGCCAGGCGTGATGAAGGACTTCATGAAGCA
GGTGCAGACCAGCGGCGTGCTGATGGTCTTCTCTCAGGCCTGGATTGAGGAGCTCTACCATCAGGTGCTCGACAGGA
ACATGCTTGGAGAGGCTGGCTATTGGGGCAGCCCAGAAGATAACAGCCTTCCCCTCATCACAATGCTGACCGATATT
GACGGCTTAGAGAGCAGTGCCATTGGTGGCCAGCTGATGGCCTCGGCTGCTACAGAGTCTCCTTTCGCCCAGAGCAG
GAGAATTGATGACTCCACAGTGGCAGGCGTGGCATTTGCTCGCTATATTCTGGTGGGGCTGCTGGAAGAACTTGATCG
ATACTTTATCAACCCCACTGACTGGTCGAATGGCGGGGAGCTCCAAAGGGCTGGCCTTCATTCTGGGGAGCTGAAGGC
ATCAAAGAGCAGAACCAGAAGGAGCGGGACGCCATCTGCATGAGCCTCGACGGGCTGCGGGAAAGCCGCCACGGCTGAG
CTGCGCTCTAGGCGTTGCTGCTAACTGCGCCTCAGCCCTTGCCCAGATGGCAGCTGCCTCCTGTGTCCAAGAAGAAA
AAGAAGAGAGGGAGGCCCAAGAACCCAGTGATGCCCATCACACAAGTGAAACTAAAAGTGGAGCAGAAACTGGAGCAG
ATTGGGAAGGTGCAGGGGGTGTGGCTGCACACTGCCCACGTCTTGTGCATGGAGGCCATCCTCAGCGTAGGCCTGGA
GATGGGAAGCCACAACCCGGACTGCTGGCCACACATGTGTTCAGGGTGTGTGAATACGTGGGCACCCTGGAGCACAACC
ACTTCAGCGATGGTGCCTCGCAGCCCCCTCTGACCATCAGCCAGCCCCACGAAGGCCACTGGAAGCGCTGGCCTCCTT
GGGGACCCCGAGTGTGAGGGCTCGCCCCCCGAGCACAGCCCGGAGCAGGGGCGCTCCCTGAGCACGGCCCCCTGTCGT
CCAGCCCCTGTCCATCCAGGACCTCGTCCGGGAAGGCAGCCGGGGTCGGGCCTCCGACTTCCGCGGCGGGAGCCTCA
TGAGCGGGAGCAGCCGGGCCAAGGTGGTGCTCACCCTCTCCACCCTCCACGAACCGACAGGCCTCTTTGAAGATGCTACGGAT
AAGTTGAACCTCATGGCCTTGGGAGGTTTTCTTTACCAGCTGAAGAAAGCATCGCAGTCTCAGCTTTTCCATTCTGT
TACAGATACAGTTGATTACTCTCTGGCAATGCCAGGAGAAGTTAAATCCACTCAAGACCGAAAAAGCGCCCCTCCACC
TGTTCCGCCTGGGGAATGCCCATGCTGAGGATTGTGCGGAGCAAAGCACGGCCCCTGCTCCACGTGATGCGCTGCTGG
AGCCTTGTGGCCCCACACCTGGTGGAGGCTGCTTGCCATAAGGAAAGACATGTGTCTCAGAAGGCTGTTTCCTTCAT
CCATGACATACTGACAGAAGTCCTCACTGACTGGAATGAGCCACCTCATTTTCACTTCAATGAAGCACTCTTCCGAC
CTTTCGAGCCGCATTATGCAGCTGGAATTGTGTGATGAGGACGTCCAAGACCAGGTTGTCACATCCATTGGTGAGCTG
GTTGAAGTGTGTTCCACGCAGATCCAGTCGGGATGGAGACCCTTGTTCAGTGCCCTGGAAACAGTGCATGGCGGGAA
CAAGTCAGAGATGAAGGAGTACCTGGTTGGTGACTACTCCATGGGAAAAGGCCAAGCCTCCAGTGTTTGATGTATTTG
AAGCTTTTCTCAATACTGACAACATCCAGGTCTTTGCTAATGCAGCCACTAGCTACATCATGTGCCTTATGAAGTTT
GTCAAAGGACTGGGGGAGGTGGACTGTATAAAGAGATTGGAGACTGTGCCCCAGCACCCGGAGCCCCGTCCACAGACCT
GTGCCTCCCGGCCCTGGATTACCTCAGGCGCTGCTCTCAGTTATTGGCCAAAATCTACAAAATGCCCTTGAAGCCAA
TATTCCTTAGTGGGAGACTTGCCGGCTTGCCTCGAAGACTTCAGGAACAGTCAGCCAGCAGTGAGGATGGAATTGAA
TCAGTCCTGTCTGATTTTGATGATGACACCGGTCTGATAGAAGTCTGGATAATCCTGCTGGAGCAGCTGACAGCGGC
TGTGTCCAATTGTCCACGGCAGCACCAACCACCAACTCTGGATTTACTCTTTGAGCTGTTGAGAGATGTGACGAAAA
CACCAGGACCAGGGTTTGGTATCTATGCAGTGGTTCACCTCCTCCTTCCTGTGATGTCCGTTTGGCTCC

FIGURE 44B

```
GCCGGAGCCATAAAGACCATTCCTACTGGGATATGGCCTCTGCCAATTTCAAGCACGCTATTGGTCTGTCCTGTGAG
CTGGTGGTGGAGCACATTCAAAGCTTTCTACATTCAGATATCAGGTACGAGAGCATGATCAATACCATGCTGAAGGA
CCTCTTTGAGTTGCTGGTCGCCTGTGTGGCCAAGCCCACTGAAACCATCTCCAGAGTGGGCTGCTCCTGTATTAGAT
ACGTCCTTGTGACAGCGGGCCCTGTGTTCACTGAGGAGATGTGGAGGCTTGCCTGCTGTGCCCTGCAAGATGCGTTC
TCTGCCACACTCAAGCCAGTGAAGGACCTGCTGGGCTGCTTCCACAGCGGCACGGAGAGCTTCAGCGGGGAAGGCTG
CCAGGTGCGAGTGGCGGCCCCGTCCTCCTCCCAAGTGCCGAGGCCGAGTACTGGCGCATCCGAGCCATGGCCCAGC
AGGTGTTTATGCTGGACACCCAGTGCTCACCAAAGACACCAAACAACTTTGACCACGCTCAGTCCTGCCAGCTCATT
ATTGAGCTGCCTCCTGATGAAAAACCAAATGGACACACCAAGAAAAGCGTGTCTTTCAGGGAAATTGTGGTGAGCCT
GCTGTCTCATCAGGTGTTACTCCAGAACTTATATGACATCTTGTTAGAAGAGTTTGTCAAAGGCCCCTCTCCTGGAG
AGGAAAAGACGATACAAGTGCCAGAAGCCAAGCTGGCTGGCTTCCTCAGATACATCTCTATGCAGAACTTGGCAGTC
ATATTCGACCTGCTGCTGGACTCTTATAGGACTGCCAGGGAGTTTGACACCAGCCCCGGGCTGAAGTGCCTGCTGAA
GAAAGTGTCTGGCATCGGGGGCGCCGCCAACCTCTACCGCCAGTCTGCGATGAGCTTTAACATTTATTTCCACGCCC
TGGTGTGTGCTGTTCTCACCAATCAAGAAACCATCACGGCCGAGCAAGTGAAGAAGGTCCTTTTTGAGGACGACGAG
AGAAGCACGGATTCTTCCCAGCAGTGTTCATCTGAGGATGAAGACATCTTTGAGGAAACCGCCCAGGTCAGCCCCCC
GAGAGGCAAGGAGAAGAGACAGTGGCGGGCACGGATGCCCTTGCTCAGCGTCCAGCCTGTCAGCAACGCAGATTGGG
TGTGGCTGGTCAAGAGGCTGCACAAGCTGTGCATGGAACTGTGCACAACTACATCCAGATGCACTTGGACCTGGAG
AACTGTATGGAGGAGCCTCCCATCTTCAAGGGCGACCCGTTCTTCATCCTGCCCTCCTTCCAGTCCGAGTCATCCAC
CCCATCCACCGGGGGCTTCTCTGGGAAAGAAACCCCTTCCGAGGATGACAGAAGCCAGTCCCGGGAGCACATGGGCG
AGTCCCTGAGCCTGAAGGCCGGTGGTGGGGACCTGCTGCTGCCCCCCAGCCCCAAAGTGGAGAAGAAGGATCCCAGC
CGGAAGAAGGAGTGGTGGGAGAATGCGGGGAACAAAATCTACACCATGGCAGCCGACAAGACCATTTCAAAGTTGAT
GACCGAATACAAAAAGAGGAAACAGCAGCACAACCTGTCCGCGTTCCCCAAAGAGGTCAAAGTGGAGAAGAAAGGAG
AGCCACTGGGTCCCAGGGGCCAGGACTCCCCGCTGCTTCAGCGTCCCCAGCACTTGATGGACCAAGGGCAAATGCGG
CATTCCTTCAGCGCAGGCCCCGAGCTGCTGCGACAGGACAAGAGGCCCCGCTCAGGCTCCACCGGGAGCTCCCTCAG
TGTCTCGGTGAGAGACGCAGAAGCACAGATCCAGGCATGGACCAACATGGTGCTAACAGTTCTCAATCAGATTCAGA
TTCTCCCAGACCAGACCTTCACGGCCCTCCAGCCCGCAGTGTTCCCCGTGCATCAGTCAGCTGACCTGTCACGTGACC
GACATCAGAGTTCGCCAGGCTGTGAGGGAGTGGCTGGGCAGGGTGGGCCGTGTCTATGCATCATTGT**TAG**CCGAC
TCCTGTTCTACTCTCCCACCAAATAACAGTAGTGAGGGTTAGAGTCCTGCCAATACAGCTGTTGCATTTTCCCCACC
ACTAGCCCCACTTAAACTACTACTACTGTCTCAGAGAACAGTGTTTCCTAATGTAAAAAGCCTTTCCAACCACTGAT
CAGCATTGGGGCCATACTAAGGTTTGTATCTAGATGACACAAACGATATTCTGATTTTGCACATTATTATAGAAGAA
TCTATAATCCTTGATATGTTTCTAACTCTTGAAGTATATTTCCCAGTGCTTTTGCTTACAGTGTTGTCCCCAAATGG
GTCATTTTCAAGGATTACTCATTTGAAAACACTATATTGATCCATTTGATCCATCATTTAAAAAATAAATACAATTC
CTAAGGCAATATCTGCTGGTAAGTCAAGCTGATAAACACTCAGACATCTAGTACCAGGGATTATTAATTGGAGGAAG
ATTTATGGTTATGGGTCTGGCTGGGAAGAAGACAACTATAAATACATATTCTTGGGTGTCATAATCAAGAAAGAGGT
GACTTCTGTTGTAAAATAATCCAGAACACTTCAAAATTATTCCTAAATCATTAAGATTTTCAGGTATTCACCAATTT
CCCCATGTAAGGTACTGTGTTGTACCTTTATTTCTGTATTTCTAAAAGAAGAAAGTTCTTTCCTAGCAGGGTTTGAA
GTCTGTGGCTTATCAGCCTGTGACACAGAGTACCCAGTGAAAGTGGCTGGTACGTAGATTGTCAAGAGACATAAGAC
CGACCAGCCACCCTGGCTGTTCTTGTGGTGTTTGTTTCCATCCCCAAGGCAAACAAGGAAAGGAAAGGAAAGAAGAA
AAGGTGCCTTAGTCCTTTGTTGCACTTCCATTTCCATGCCCCACAATTGTCTGAACATAAGGTATAGCATTTGGTTT
TTAAGAAAACAAAACATTAAGACGCAACTCATTTTATATCAACACGCTTGGAGGAAAGGGACTCAGGGAAGGGAGCA
GGGAGTGTGGGGTGGGGATGGATTATGATGAAATCATTTTCAATCTTAAAATATAATACAACAATCTTGCAAAATTA
TGGTGTCAGTTACACAAGCTCTAGTCTCAAAATGAAAGTAATGGAGAAAAGACACTGAAATTTAGAAAATTTTGTCGA
TTTAAAATATTTCTCCTATCTACCAAGTAAAGTTACCCTATGTTTGATGTCTTTGCATTCAGACCAATATTTCAGGT
GGATATTTCTAAGTATTACTAGAAAATACGTTTGAAAGCTTTATCTTATTATTTACAGTATTTTTATATTTCTTACA
TTATCCTAATGATTGAAAACTCCTCAATCAAGCTTACTTACACACATTCTACAGAGTTATTTAAGGCATACATTATA
ATCTCCCAGCCCCATTCATAATGAATAAGTGACCCCTTTAAATATAAGACACAAATTCTACAGTATTGAAATAAGGAT
TTAAAGGGGTATTTGTAAACTTTGCCCTCCTTGAGAAATATGGAACTACCTTAGAGGTTAAGAGGAAGGCAGTGTTC
TGACTTCTTTAGGTGATCTGAAAAAAAACACCCTTTATCATCCAGTGTACCATCTAGAGATCACCACAGAATCCATTTT
TTTCCCAGTTCCACAAAACACTCTGTTTGCCTTCAGTTTTTACTCACTAGACAATAATTCAAGTTTAGAAACAGGTA
ATCAGCTATTTGATCTTAAAAGGCAATGAATTGTTGGGATATCAGTGAACTATGTTGTATACTTTTGAATTTTTACA
TTTTATAAATGGAATTGAAAGTTGGATAACTGCTTTTTTTAAATTTTCCAACAGAAGTAACACCACAGTTGCTTT
```

FIGURE 44C

GTTTCTTTTTATAGCTTACCTGAGGTTCAGTTCTTCTTTGTGAACCTGTGAGTACTCCACAGTTTACTGGGGAAAA
GGCTTCAGTAAAGCAGAGGCTAGAATTACAGTATTTATACATAGCAACTTTTCATAAAGTAGAAAAATTCAAAGGAA
GCTGTCTCAATTTGAGAATACCAGCTGGGCACGGTCG

FIGURE 45

```
CAGAGAGGCTGTATTTCAGTGCAGCCTGCCAGACCTCTTCTGGAGGAAGACTGGACAAAGGGGGTCACACATTCCTT
CCATACGGTTGAGCCTCTACCTGCCTGGTGCTGGTCACAGTTCAGCTTCTTCATGATGGTGGATCCCAATGGCAATG
AATCCAGTGCTACATACTTCATCCTAATAGGCCTCCCTGGTTTAGAAGAGGCTCAGTTCTGGTTGGCCTTCCCATTG
TGCTCCCTCTACCTTATTGCTGTGCTAGGTAACTTGACAATCATCTACATTGTGCGGACTGAGCACAGCCTGCATGA
GCCCATGTATATATTTCTTTGCATGCTTTCAGGCATTGACATCCTCATCTCCACCTCATCCATGCCCAAAATGCTGG
CCATCTTCTGGTTCAATTCCATACCATCCAGTTTGATGCTTGTCTGCTACAGATTTTTGCCATCCACTCCTTATCT
GGCATGGAATCCACAGTGCTGCTGGCCATGGCTTTTGACCGCTATGTGGCCATCTGTCACCCACTGCGCCATGCCAC
AGTACTTACGTTGCCTCGTGTCACCAAAATTGGTGTGGCTGCTGTGGTGCGGGGGCTGCACTGATGGCACCCCTTC
CTGTCTTCATCAAGCAGCTGCCCTTCTGCCGCTCCAATATCCTTTCCCATTCCTACTGCCTACACCAAGATGTCATG
AAGCTGGCCTGTGATGATATCCGGGTCAATGTCGTCTATGGCCTTATCGTCATCATCTCCGCCATTGGCCTGGACTC
ACTTCTCATCTCCTTCTCATATCTGCTTATTCTTAAGACTGTGTTGGGCTTGACACGTGAAGCCCAGGCCAAGGCAT
TTGGCACTTGCGTCTCTCATGTGTGTGCTGTGTTCATATTCTATGTACCTTTCATTGGATTGTCATGGTGCATCGC
TTTAGCAAGCGGCGTGACTCTCCGCTGCCCGTCATCTTGGCCAATATCTATCTGCTGGTTCCTCCTGTGCTCAACCC
AATTGTCTATGGAGTGAAGACAAAGGAGATTCGACAGCGCATCCTTCGACTTTTCCATGTGGCCACACACGCTTCAG
AGCCCTAGGTGTCAGTGATCAAACTTCTTTTCCATTCAGAGTCCTCTGATTCAGATTTTAATGTTAACATTTTGGAA
GACAGTATTCAGAAAAAAATTTCCTTAATAAAAAATACAACTCAGATCCTTCAAATATGAAACTGGTTGGGGAATC
TCCATTTTTTCAATATTATTTTCTTCTTTGTTTTCTTGCTACATATAATTATTAATACCCTGACTAGGTTGTGGTTG
GAGGGTTATTACTTTTCATTTTACCATGCAGTCCAAATCTAAACTGCTTCTACTGATGGTTTACAGCATTCTGAGAT
AAGAATGGTACATCTAGAGAACATTTGCCAAAGGCCTAAGCACGGCAAAGGAAAATAAACACAGAATATAATAAAAT
GAGATAATCTAGCTTAAAACTATAACTTCCTCTTCAGAACTCCCAACCACATTGGATCTCAGAAAAATGCTGTCTTC
AAAATGACTTCTACAGAGAAGAAATAATTTTTCCTCTGGACACTAGCACTTAAGGGGAAGATTGGAAGTAAAGCCTT
GAAAAGAGTACATTTACCTACGTTAATGAAAGTTGACACACTGTTCTGAGAGTTTTCACAGCATATGGACCCTGTTT
TTCCTATTTAATTTTCTTATCAACCCTTTAATTAGGCAAAGATATTATTAGTACCCTCATTGTAGCCATGGGAAAAT
TGATGTTCAGTGGGGATCAGTGAATTAAATGGGGTCATACAAGTATAAAAATTAAAAAAAAAAAAGACTTCATGCCC
AATCTCATATGATGTGGAAGAACTGTTAGAGAGACCAACAGGGTAGTGGGTTAGAGATTTCCAGAGTCTTACATTTT
CTAGAGGAGGTATTTAATTTCTTCTCACTCATCCAGTGTTGTATTTAGGAATTTCCTGGCAACAGAACTCATGGCTT
TAATCCCACTAGCTATTGCTTATTGTCCTGGTCCAATTGCCAATTACCTGTGTCTTGGAAGAAGTGATTTCTAGGTT
CACCATTATGGAAGATTCTTATTCAGAAAGTCTGCATAGGGCTTATAGCAAGTTATTTATTTTTAAAAGTTCCATAG
GTGATTCTGATAGGCAGTGAGGTTAGGGAGCCACCAGTTATGATGGGAAGTATGGAATGGCAGGTCTTGAAGATAAC
ATTGGCCTTTTGAGTGTGACTCGTAGCTGGAAAGTGAGGGAATCTTCAGGACCATGCTTTATTTGGGGCTTTGTGCA
GTATGGAACAGGGACTTTGAGACCAGGAAAGCAATCTGACTTAGGCATGGGAATCAGGCATTTTTGCTTCTGAGGGG
CTATTACCAAGGGTTAATAGGTTTCATCTTCAACAGGATATGACAACAGTGTTAACCAAGAAACTCAAATTACAAAT
ACTAAAACATGTGATCATATATGTGGTAAGTTTCATTTTCTTTTTCAATCCTCAGGTTCCCTGATATGGATTCCTAT
AACATGCTTTCATCCCCTTTTGTAATGGATATCATATTTGGAAATGCCTATTTAATACTTGTATTTGCTGCTGGACT
GTAAGCCCATGAGGGCACTGTTTATTATTGAATGTCATCTCTGTTCATCATTGACTGCTCTTTGCTCATCATTGAAT
CCCCCAGCAAAGTGCCTAGAACATAATAGTGCTTATGCTTGACACCGGTTATTTTTCATCAAACCTGATTCCTTCTG
TCCTGAACACATAGCCAGGCAATTTTCCAGCCTTCTTTGAGTTGGGTATTATTAAATTCTGGCCATTACTTCCAATG
TGAGTGGAAGTGACATGTGCAATTTCTATACCTGGCTCATAAAACCCTCCCATGTGCAGCCTTTCATGTTGACATTA
AATGTGACTTGGGAAGCTATGTGTTACACAGAGTAAATCACCAGAAGCCTGGATTTCTGAAAAAACTGTGCAGAGCC
AAACCTCTGTCATTTGCAACTCCCACTTGTATTTGTACGAGGCAGTTGGATAAGTGAAAAATAAAGTACTATTGTGT
CAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 46

```
AGCACCTTGGCACAGCTGAAGCAGAGGAACACGCTGAAGGATGGTATCATCATGATCCAGACGCTGCTGATCATCCT
CTTCATCATCGTGCCTATCTTCCTGCTGCTGGACAAGGATGACAGCAAGGCTGGCATGGAGGAAGATCACACCTACG
AGGGCCTGGACATTGACCAGACAGCCACCTATGAGGACATAGTGACGCTGCGGACAGGGGAAGTGAAGTGGTCTGTA
GGTGAGCACCCAGGCCAGGAGTGAGAGCCAGGTCGCCCCATGACCTGGGTGCAGGCTCCCTGGCCTCAGTGACTGCT
TCGGAGCTGCCTGGCTCATGGCCCAACCCCTTTCCCGGACCCCCCAGCTGGCCTCTGAAGCTGGCCCACCAGAGCTG
CCATTTGTCTCACCCCTGGTGTCCAGCTCTTGCCAAAGGGCCTGGAGTNNAAGGACAACAGGCANCACTTGGAGGGA
GTTCTCTGGGGATGGACGGGACCCACCTT
```

FIGURE 47

CAGGAAGGTTCCTCTCCCAGTGGCC**ATG**GGTAGCAACAGTGGGCAGGCTGGCCGCCACATCTATAAATCCCTAGCTG
ATGATGGCCCCTTTGACTCTGTGGAGCCGCCTAAAAGACCCACCAGCAGACTCATCATGCACAGCATGGCCATGTTC
GGAAGAGAGTTCTGCTACGCGGTGGAGGCAGCGTATGTGACCCCAGTCCTGCTCAGCGTAGGTCTGCCCAGCAGCCT
GTACAGCATTGTGTGGTTCCTCAGCCCCATCCTGGGATTCCTGCTGCAGCCCGTGGTCGGATCGGCCAGCGACCACT
GCCGGTCCAGGTGGGGCCGCCGGAGACCCTACATCCTCACCCTGGGAGTCATGATGCTCGTGGGCATGGCTCTGTAC
CTCAATGGGGCTACTGTTGTAGCAGCTTTGATTGCTAACCCAAGGAGGAAGCTGGTTTGGGCCATAAGTGTCACCAT
GATAGGTGTCGTTCTCTTTGATTTTGCTGCCGACTTCATTGATGGGCCCATCAAAGCCTACTTATTTGATGTCTGCT
CCCATCAGGACAAGGAGAAGGGCCTCCACTACCATGCCCTCTTCACAGGTTTTGGAGGTGCCCTGGGTTACCTTTTG
GGTGCTATAGACTGGGCCCATCTGGAGCTGGGAAGACTGTTGGGTACAGAATTCCAGGTCATGTTCTTCTTCTCTGC
ATTGGTGCTCACTTTGTGTTTTACTGTTCATCTGTGCAGTATCTCTGAAGCCCCACTTACAGAGGTTGCAAAGGGCA
TTCCCCCACAGCAAACCCCTCAGGACCCTCCATTGTCATCAGATGGAATGTACGAGTATGGTTCTATCGAGAAAGTT
AAAAATGGTTACGTAAATCCAGAGCTGGCAATGCAGGGAGCAAAAAACAAAAATCATGCTGAACAGACTCGCAGGGC
AATGACATTAAAGTCACTGCTGAGAGCACTGGTGAACATGCCTCCTCACTACCGCTACCTTTGCATCAGCCACCTCA
TTGGATGGACGGCCTTCCTGTCCAACATGCTGTTCTTCACAGATTTCATGGGCCAGATTGTGTACCGCGGGGATCCC
TATAGTGCACACAACTCCACAGAGTTTCTCATCTACGAAAGAGGAGTCGAGGTTGGATGTTGGGGCTTCTGCATCAA
CTCCGTGTTTTCCTCACTTTATTCTTACTTTCAGAAAGTTTTGGTATCCTACATTGGATTAAAGGGTCTTTACTTCA
CGGGATATTTGCTGTTTGGCCTGGGGACGGGATTTATTGGGCTCTTCCCGAATGTCTACTCCACCCTGGTCCTGTGC
AGCCTGTTTGGTGTAATGTCCAGCACCCTGTACACTGTGCCCTTTAACCTCATTACTGAGTACCACCGCGAGGAAGA
AAAGGAGAGGCAGCAGGCCCCAGGAGGGGACCCAGACAACAGCGTGAGAGGGAAGGGCATGGACTGCGCGCACCCTCA
CATGCATGGTGCAGCTGGCTCAGATCCTGGTCGGAGGTGGCCTGGGCTTTCTGGTCAACACAGCCGGGACCGTTGTC
GTCGTGGTGATCACAGCGTCTGCGGTGGCACTGATAGGCTGTTGCTTTGTCGCTCTCTTTGTTAGATATGTGGAT**TA**
GGTCAATAAAGAGACAATGACCCTAAAAAAAAA

FIGURE 48

```
GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCATCTTCATCATTCATAT
GAGGAAATAAGTGGTAAAATCCTTGGAAATACAATGAGACTCATCAGAAACATTTACATATTTTGTAGTATTGTTAT
GACAGCAGAGGGTGATGCTCCAGAGCTGCCAGAAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAA
AGGTTCCCGCAGACTTGACCCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCA
GATTTTCATTCTGTCTCCAAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGGATCTCAAAACCTT
TGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTGAAGAGTGTAACTTGGTATTTACTGGCAG
GTCTCAGGTATTTAGATCTTTCTTTTAATGACTTTGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTCACAC
CTGGAAATCCTAGGTTTGAGTGGGGCAAAAATACAAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATAC
TGTCTTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAACAAAACTGCACA
TTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCGTGATGGAATCAAGACTTCAAAAATATTAGAAATGACA
AATATAGATGGCAAAAGCCAATTTGTAAGTTATGAAATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGT
TCTATTGCTTAATAAAGTTGATTTACTCTGGGACGACCTTTTCCTTATCTTACAATTTGTTTGGCATACATCAGTGG
AACACTTTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTACTCAAATACT
GTAATGAGAACTATAAAATTGGAGCATGTACATTTCAGAGTGTTTTACATTCAACAGGATAAAATCTATTTGCTTTT
GACCAAAATGGACATAGAAAACCTGACAATATCAAATGCACAAATGCCACACATGCTTTTCCCGAATTATCCTACGA
AATTCCAATATTTAAATTTTGCCAATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACTGCCTCACTTG
AAAACTCTCATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACACCCTTGGA
ACACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTGCTCATGGCCAGAAACTGTGGTCAATA
TGAATCTGTCATCAATAAATTGTCTGATTCTGTCTTCAGGTGCTTGCCCAAAAGTATTCAAATACTTGACCTAAAT
AATAACCAAATCCAAACTGTACCTAAAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTT
TCTAACTGATCTCCCTGGATGCAGTCATTTCAGTAGACTTTCAGTTCTGAACATTGAAATGAACTTCATTCTCAGCC
CATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAAACTCTAAATGCGGGAAGAAATCCATTCCGGTGTACCTGT
GAATTAAAAAATTTCATTCAGCTTGAAACATATTCAGAGGTCATGATGGTTGGATGGTCAGATTCATACACCTGTGA
ATACCCTTTAAACCTAAGGGGAACTAGGTTAAAAGACGTTCATCTCCACGAATTATCTTGCAACACAGCTCTGTTGA
TTGTCACCATTGTGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCCTGG
TATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCAAGAACAACTCAAGAGAAA
TGTCCGATTCCACGCATTTATTTCATACAGTGAACATGATTCTCTGTGGGTGAAGAATGAATTGATCCCCAATCTAG
AGAAGGAAGATGGTTCTATCTTGATTTGCCTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATT
GTAAGCTTCATTGAGAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATTA
TGAATTCTACTTTGCCCACCACAATCTCTTCCATGAAAATTCTGATCATATAATTCTTATCTTACTGGAACCCATTC
CATTCTATTGCATTCCCACCAGGTATCATAAACTGAAAGCTCTCCTGGAAAAAAAAGCATACTTGGAATGGCCCAAG
GATAGGCGTAAATGTGGGCTTTTCTGGGCAAACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAAT
GTATGAACTGCAGACATTCACAGAGTTAAATGAAGAGTCTCGAGGTCTTCTCTCTGATGAGAACAGATTGTCTA
TAAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATACAACCTTTATGATG
GCAATTTGACAATATTTATTAAAATAAAAAATGGTTATTCCCTTCATATCAGTTTCTAGAAGGATTTCTAAGAATGT
ATCCTATAGAAACACCTTCACAAGTTTATAAGGGCTTATGGAAAAAGGTGTTCATCCCACGGATTGTTTATAATCATG
AAAAATGTGGCCAGGTGCAGTGGCTCACTCTTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGGT
CAAGAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGAT
GGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGTGGCAGTTGCAG
TGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGTGACAGAGCGAGACTCCATCTCAAAAAAAAGAAAAAAAAAAAA
GAAAAAAATGGAAAACATCCTCATGGCCACAAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAA
TATTACATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCCTGGTATGGAAAAAACATATTAATATGTTATAAACT
ATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAAATAAAAGTGTAAAGAAATCTATACC
AGATGTAGTAACAGTCGTTTGGGTCTGGGAGGTTGGATTACAGGGAGCATTTGATTTCTATGTTCTGTATTTCTATA
ATGTTTGAATTGTTTAGAATGAATCTGTATTTCTTTTATAAGTAGAAAAAAAAATAAAGATAGTTTTTACAGCCT
```

FIGURE 49A

```
GTCACCTGGATGGTGTAAAGGAAACAAGAGACAGGAACAGAGCCCCTCATCTCACCTCTGGGCTACCATACAGAAAA
GCAGAATTGGCAGGAACTGAAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATC
GTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGATGGCCCCTGGAGT
CAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCCTTGAGAAC
CATGATTCCCTTCCGTCCCAAGCCGAGGTTTCCTGCCCCCCAGCCCCTGGACAATGCTGGCCTGTTCTCCTACCTCA
CCGTGTCATGGCTCACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTGTCA
GTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAAGTCTCAAGGCGAGGGATTGA
AAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGAACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCT
TCTGCATTGCCAGTGTACTCGGGCCAATATTGATTATACCAAAGATCCTGGAATATTCAGAAGAGCAGTTGGGGAAT
GTTGTCCATGGAGTGGGACTCTGCTTTGCCCTTTTTCTCTCCGAATGTGTGAAGTCCTGAGTTTCTCCTCCAGTTG
GATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCCTTTGCCTTTGAGAAGCTCATCCAATTTA
AGTCTGTAATACACATCACCTCAGGAGAGGCCATCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTG
TGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCCTACTTCATTATTGGATA
CACTGCATTTATTGCCATCTTATGCTATCCCCTGGTTTTCCCACTGGAGGTATTCATGACAAGAATGGCTGTGAAGG
CTCAGCATCACACATCTGAGGTCAGCGACCAGCGCATCCGTGTGACCAGTGAAGTTCTCACTTGCATTAAGCTGATT
AAAATGTACACATGGGAGAAACCATTTGCAAAAATCATTGAAGACCTAAGAAGGAAGGAAAGGAAGCTATTGGAGAA
GTGCGGGCTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCGCAGTGGCCACAGCGGTCTGGGTTCTCA
TCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTAGCCTCCTTGAATCTCCTTCGGCTG
TCAGTGTTCTTTGTGCCTATTGCAGTCAAAGGTCTCACGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTT
CCTCCAGGAGAGCCCTGTTTTCTATGTCCAGACATTACAAGACCCCAGCAAAGCTCTGGTCTTTGAGGAGGCCACCT
TGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGG
ATGACCAGGCCTAGAGATGCCCTCGGGCCAGAGGAAGAAGGGAACAGCCTGGGCCCAGAGTTGCACAAGATCAACCT
GGTGGTGTCCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTCAGCCATCC
TGGAGGAGATGCACTTGCTCGAGGGCTCGGTGGGGGTGCAGGGAAGCCTGGCCTATGTCCCCCAGCAGGCCTGGATC
GTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCCGATACCTCCAGGTGCTCCACTG
CTGCTCCCTGAATCGGGACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGGCCCAACCTCT
CTGGGGGGCAGAAACAGAGGATCAGCCTGGCCCGCGCCGTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC
CTGTCTGCTGTGGACGCCCACGTGGGGAAGCACATTTTTGAGGAGTGCATTAAGAAGACACTCAGGGGGAAGACGGT
CGTCCAGGTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGCCAGGTCATTTTGTTGGAAAATGGGAAAATCTGTG
AAAATGGAACTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCC
ACTTCGGACATGTTGCAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCT
GGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAGGAGGGAGATGGAAGAAGGCTCCTTGA
GTTGGAGGGTCTACCACCACTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCATAATTTTCTTCTTTGTGGTG
CTGATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAG
CCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCCTTCTACCAGCTGGTGT
ACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAGGTCACGAGGAAGGCA
TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTCCGCTGCCCCATGAGTTTCTTTGACACCATCCCAATAGG
CCGGCTTTTGAACTGCTTCGCAGGGGACTTGGAACAGCTGGCACAGCTCTTGCCCATCTTTTCAGAGCAGTTCCTGG
TCCTGTCCTTAATGGTGATCGCCGTCCTGTTGATTGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATA
ATCATGGTTATTTGCTTCATTTATTATATGATGTTCAAGGAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAG
CCGGTCTCCTTTAATGGTGATCGCCGTCCTGTTGATTGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATA
```

FIGURE 49B

```
CTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGACACCCTGATCCAGCGCACAATCCGTGAAGCCTTCCA
GGGCTGCACCGTGCTCGTCATTGCCCACCGTGTCACCACTGTGCTGAACTGTGACCGCATCCTGGTTATGGGCAATG
GGAAGGTGGTAGAATTTGATCGGCCGGAGGTACTGCGGAAGAAGCCTGGGTCATTGTTCGCAGCCCTCATGGCCACA
GCCACTTCTTCACTGAGA**TAA**GGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCACACAGCTG
CAGCTTCGAGGCCCACAGTCTGCGACCTTCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGG
GGGGTGGGGATTGCTGGATGGAAACCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAGAACCATC
TAAGACATGGGATTCAGTGATCATGTGGTTCTCC
```

FIGURE 50

```
GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAGGA
CAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCCTGCCGTCAGCTGCCGGCCGAGTT
GGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGGAGATTG
TCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGATTTCAAAGC
TGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATC**ATG**TCTGTGAT
GGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGGAACACCTTTTGCTGTGATGGCCGCGTCA
TGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCCTTTTCCTCATCCTGGGGACATGTACACTCTTCTTCGCCTTT
GAGTGCCGCTACCTGGCTGTTCAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCTTCCTTTTCTCCATGGC
TACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCCTCGGGCGCTACCAGATGAAGCAGCTTTCATAGAAA
TGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTATCAAGAATTTCCAGATAAAC
AACCAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCCGGGCCTCCCATTGCAGCATCTG
TGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTACCGCTACT
TCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCTTCAACATCGTCTATGTGGCCCTCAAA
TCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAACTGTTCTAGAAGTCCTCATTTGCTTCTTTAC
ACTCTGGTCCGTCGTGGGACTGACTGGATTTCATACTTTCCTCGTGGCTCTCAACCAGACAACCAATGAAGACATCA
AAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCCATGGCAATATTGTGAAGAACTGCTGTGAAGTG
CTGTGTGGCCCCTTGCCCCCCAGTGTGCTGGATCGAAGGGGTATTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCC
CAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCCCAGCCCCCACAGAACACCTGAACTCAAATGAGATGC
CGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAGAGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAG
AAG**TAG**CCTATCTATGGAAGAGACTTTTGTTTGTGTTTAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACC
TGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACTGTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATT
TTCTTGCTGCAAGCTTTTTTAAATTTCTGAACTCAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAA
AATGGGTCTCTTGGGCCCTGGCACTGGTTCTCCATGGCCTCAGCCACAGGGTCCCCTTGGACCCCCTCTCTTCCCTC
CAGATCCCAGCCCTCCTGCTTGGGGTCACTGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCC
CAAGCTGCTGCACGTGCTGAGTCCAGAGGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTTCTTGGGGT
CTTCAGGACTGAAGAGGAGGGAGAGTGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCACAAATCC
TTTTAGGAATGGGACAGGTACCTTCCACTTGTTGTAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTGTTTTTCCTTTT
GACTCCTGCTCCCATTAGGAGCAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCCTCAGAGGAAG
CCCGAGTGCTCACTTAAACACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGA
AACCAAGGCACAGAGAGGCTCTCCTCTCCTCTCCTCTCCCCCGATGTACCCTCAAAAAAAAAAAAATGCTAACCAGT
TCTTCCATTAAGCCTCGGCTGAGTGAGGGAAAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCCTAAGGCCTCGG
CCCACCTCTGGCTATGGTAACCACACTGGGGGCTTCCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCC
CAGAGCCACTTCACCCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTCAGGGAAGAA
GATTTATGTATTATATGTGGCTATATTTCCTAGAGCACCTGTGTTTTCCTCTTTCTAAGCCAGGGTCCTGTCTGGAT
GACTTATGCGGTGGGGGAGTGTAAACCGGAACTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA
```

FIGURE 51

```
GAGCCGCCGCCGCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCCGGCCCCCCACCCACGTCTGCGTTGCTGCCCCGC
CTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCTCCGCCGGAGTCGAATTTACGTGCAGCTGCC
GGCAACCACAGGTTCCAAGATGGTTTGCGGGGGCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTT
ACACCTTGGTTAGTCTGCTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTG
GTCGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCCTGATTGCTTTAGTGGGTCTGATTGGAGCTGTAAAACATCA
TCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATTGTTCAGTTTTCTGTATCTTGCGCTTGTT
TAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATC
CAGAGAAATCTAAACTGCTGTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGA
CCACTCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGGTGGCATTGGCC
TGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACAGGAACCAGAAAGACCCCGCGCGAAT
CCTAGTGCATTCCTTTGATGAGAAAACAAGGAAGATTTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTC
TGTTAAGCTCCATTTGCCAGTTTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATT
TTTACTCTATGTTTCTCTACATGTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTCTGAAGCTCG
GTGGCACCTGGAATTTACTGTATTCATTGTCGGGCACTGTCCACTGTGGCCTTTCTTAGCATTTTTACCTGCAGAAA
AACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAATCTGAACGTACATCTCACTGGTATAATTATATGTAGCAC
TGTGCTGTGTAGATAGTTCCTACTGCAAAAAGAGTGGAAATTTATTAAAATCAGAAAGTATGAGATCCTGTTATGTT
AAGGGAAATCCAAATTCCCAATTTTTTTTGGTCTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTATAAAAAT
GATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTATGTCTTAGGAAATTGTGGTTTAA
TTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCA
GCCTCCATCAGAATGGAACGAGTTTTGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATT
TGAAGTCTAAAAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGATATTTGATTAT
CTTAAAAATTGTTAAATACCGTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTTGTCAAACCTAAGCATATTTG
AATATGATCTCCCATAATTTGAAATTGAAATCGTATTGTGTGGCTCTGTATATTCTGTTAAAAAAATTAAAGGACAGA
AACCTTTCTTTGTGTATGCATGTTTGAATTAAAAGAAAGTAATGGAAG
```

FIGURE 52

```
CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTGCTCTGGGTGGCCCAG
ATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGTCTGCACTGAGGAGAGCAGCTGCCACACGGA
GGATGACTTGACTGATGCAAGGGAAGCTGGCTTCCAGGTCAAGGCCTACACTTTCAGTGAACCCTTCCACCTGATTG
TGTCCTATGACTGGCTGATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAG
GCCTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCCCCGGGCCTAACAG
GGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCACTGCAGTGGCATCTTCCAGAGCCCTGGTC
CTGGGATCCCAGAAACAGCATCTGTTGTGGCTATCACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTA
CCCTCAGCTGAACCCCAAGCAGGAAGCCCCATGACCCTGAGTTGTCAGACAAAGTTGCCCCTGCAGAGGTCAGCTGC
CCGCCTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATTCCAGATCCCCA
CAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGGACAACCAAGTTTGGAAACAGAGCCCC
CAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCTGCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGC
TGCTCCAGGAACTGCTCCTGAGGAGGCCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCT
TTTCTTCTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATGCAGGATGTG
AGAGTCCTCCTCGGTCACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCAGAAGCCTGGGACCACAAAGGCTAC
TGCTGAATAGAAGTAAACAGTTCATCCATGATCTCACTTAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCC
TGTCCTGCACATATGCATAAGTACTTTTACAAGTTGTCCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAA
ATAAATTTATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTCTGCTGTC
TAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAACTAATGGAAGTGGATTGAATAC
AGCAGTCTCAACTGGGGGCAATTTTGCCCCCCAGAGGACATTGGGCAATGTTTGGAGACATTTTGGTCATTATACTT
GGGGGGTTGGGGGATGGTGGGATGTGTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTA
TAATGCACAGGGCAGTACCCCACAACGAAAAATAATCTGGCCCAAAATGTCAGTTGTACTGAGTTTGAGAAACCCCA
GCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTATTATCTCTTTCCAGCCTCATT
CAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGCTATGGTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAA
AGCCATTATGTTGAAATCCTAATCCCCAAGGTGATGGCATTAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAG
TGCAGAGCCCTCATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCCTTCCACCATATG
AGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTCGCCAAACACCGACTCTGTCGTTGCCTTGA
TCTTGAACTTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGGTTGTTTGTAGCCTAA
```

FIGURE 53

CCGGCTCGCGCCCTCCGGGCCCAGCCTCCCGAGCCTTCGGAGCGGGCGCCGTCCCAGCCCAGCTCCGGGGAAACGCG
AGCCGCG**ATG**CCTGGGGGGTGCTCCCGGGGCCCCGCCGCCGGGGACGGGCGTCTGCGGCTGGCGCGACTAGCGCTGG
TACTCCTGGGCTGGGTCTCCTCGTCTTCTCCCACCTCCTCGGCATCCTCCTTCTCCTCCTCGGCGCCGTTCCTGGCT
TCCGCCGTGTCCGCCCAGCCCCCGCTGCCGGACCAGTGCCCCGCGCTGTGCGAGTGCTCCGAGGCAGCGCGCACAGT
CAAGTGCGTTAACCGCAATCTGACCGAGGTGCCCACGGACCTGCCCGCCTACGTGCGCAACCTCTTCCTTACCGGCA
ACCAGCTGGCCGTGCTCCCTGCCGGCGCCTTCGCCCGCCGGCCGCCGCTGGCGGAGCTGGCCGCGCTCAACCTCAGC
GGCAGCCGCCTGGACGAGGTGCGCGCGGGCGCCTTCGAGCATCTGCCCAGCCTGCCGCCAGCTCGACCTCAGCCACAA
CCCCACTGGCCGACCTCAGTCCCTTCGCTTTCTCGGGCAGCAATGCCAGCGTCTCGGCCCCCAGTCCCCTTGTGGAAC
TGATCCTGAACCACATCGTGCCCCCTGAAGATGAGCGGCAGAACCGGAGCTTCGAGGGCATGGTGGTGGCGGCCCTG
CTGGCGGGCCGTGCACTGCAGGGGCTCCGCCGCTTGGAGCTGGCCAGCAACCACTTCCTTTACCTGCCGCGGGATGT
GCTGGCCCAACTGCCCAGCCTCAGGCACCTGGACTTAAGTAATAATTCGCTGGTGAGCCTGACCTACGTGTCCTTCC
GCAACCTGACACATCTAGAAAGCCTCCACCTGGAGGACAATGCCCTCAAGGTCCTTCACAATGGCACCCTGGCTGAG
TTGCAAGGTCTACCCCACATTAGGGTTTTCCTGGACAACAATCCCTGGGTCTGCGACTGCCACATGGCAGACATGGT
GACCTGGCTCAAGGAAACAGAGGTAGTGCAGGGCAAAGACCGGCTCACCTGTGCATATCCGGAAAAAATGAGGAATC
GGGTCCTCTTGGAACTCAACAGTGCTGACCTGGACTGTGACCCGATTCTTCCCCCATCCCTGCAAACCTCTTATGTC
TTCCTGGGTATTGTTTTAGCCCTGATAGGCGCTATTTTCCTCCTGGTTTTGTATTTGAACCGCAAGGGGATAAAAAA
GTGGATGCATAACATCAGAGATGCCTGCAGGGATCACATGGAAGGGTATCATTACAGATATGAAATCAATGCGGACC
CCAGATTAACAAACCTCAGTTCTAACTCGGATGTC**TGA**GAAATATTAGAGGACAGACCAAGGACAACTCTGCATGAG
ATGTAGACTTAAGCTTTATCCCTACTAGGCTTGCTCCACTTTCATCCTCCACTATAGATACAACGGACTTTGACTAA
AAGCAGTGAAGGGGATTTGCTTCCTTGTTATGTAAAGTTTCTCGGTGTGTTCTGTTAATGTAAGACGATGAACAGTT
GTGTATAGTGTTTTACCCTCTTCTTTTTCTTGGAACTCCTCAACACGTATGGAGGGATTTTTCAGGTTTCAGCATGA
ACATGGGCTTCTTGCTGTCTGTCTCTCTCTCAGTACAGTTCAAGGTGTAGCAAGTGTACCCACACAGATAGCATTCA
ACAAAAGCTGCCTCAACTTTTTCGAGAAAAATACTTTATTCATAAATATCAGTTTTATTCTCATGTACCTAAGTTGT
GGAGAAAATAATTGCATCCTATAAACTGCCTGCAGACGTTAGCAGGCTCTTCAAAATAACTCCATGGTGCACAGGAG
CACCTGCATCCAAGAGCATGCTTACATTTTACTGTTCTGCATATTACAAAAAATAACTTGCAACTTCATAACTTCTT
TGACAAAGTAAATTACTTTTTTGATTGCAGTTTATATGAAAATGTACTGATTTTTTTTTAATAAACTGCATCGACAT
CCAACCGACTGAATTGTTAAAAAAAAAAAAAAAAATAAAGATTCTTAAAAGAA

FIGURE 54

```
CGGCGAGCGAGCACCTTCGACGCGGTCCGGGGACCCCCTCGTCGCTGTCCTCCCGACGCGGACCCGCGTGCCCCAGG
CCTCGCGCTGCCCGGCCGGCTCCTCGTGTCCCACTCCCGGCGCACGCCCTCCCGCGAGTCCCGGGCCCCTCCCGCGC
CCCTCTTCTCGGCGCGCGCGCAGCATGGCGCCCCCGCAGGTCCTCGCGTTCGGGCTTCTGCTTGCCGCGGCGACGGC
GACTTTTGCCGCAGCTCAGGAAGAACTGCGTGTCTGTGAAAACTACAAGCTGGCCGTAAACTGCTTTGTGAATAATAATC
GTCAATGCCAGTGTACTTCAGTTGGTGCACAAAATACTGTCATTTGCTCAAAGCTGGCTGCCAAATGTTTGGTGATG
AAGGCAGAAATGAATGGCTCAAAACTTGGGAGAAGAGCAAAACCTGAAGGGGCCCTCCAGAACAATGATGGGCTTTA
TGATCCTGACTGCGATGAGAGCGGGCTCTTTAAGGCCAAGCAGTGCAACGGCACCTCCACGTGCTGGTGTGTGAACA
CTGCTGGGGTCAGAAGAACAGACAAGGACACTGAAATAACCTGCTCTGAGCGAGTGAGAACCTACTGGATCATCATT
GAACTAAAACACAAAGCAAGAGAAAAACCTTATGATAGTAAAAGTTTGCGGACTGCACTTCAGAAGGAGATCACAAC
GCGTTATCAACTGGATCCAAAATTTATCACGAGTATTTTGTATGAGAATAATGTTATCACTATTGATCTGGTTCAAA
ATTCTTCTCAAAAAACTCAGAATGATGTGGACATAGCTGATGTGGCTTATTATTTTGAAAAAGATGTTAAAGGTGAA
TCCTTGTTTCATTCTAAGAAAATGGACCTGACAGTAAATGGGGAACAACTGGATCTGGATCCTGGTCAAACTTTAAT
TTATTATGTTGATGAAAAAGCACCTGAATTCTCAATGCAGGGTCTAAAAGCTGGTGTTATTGCTGTTATTGTGGTTG
TGGTGATAGCAGTTGTTGCTGGAATTGTTGTGCTGGTTATTTCCAGAAAGAAGAGAATGGCAAAGTATGAGAAGGCT
GAGATAAAGGAGATGGGTGAGATGCATAGGGAACTCAATGCATAACTATATAATTTGAAGATTATAGAAGAAGGGAA
ATAGCAAATGGACACAAATTACAAATGTGTGTGCGTGGGACGAAGACATCTTTGAAGGTCATGAGTTTGTTAGTTTA
ACATCATATATTTGTAATAGTGAAACCTGTACTCAAAATATAAGCAGCTTGAAACTGGCTTTACCAATCTTGAAATT
TGACCACAAGTGTCTTATATATGCAGATCTAATGTAAAATCCAGAACTTGGACTCCATCGTTAAAATTATTTATGTG
TAACATTCAAATGTGTGCATTAAATATGCTTCCACAGTAAAATCTGAAAAACTGATTTGTGATTGAAAGCTGCCTTT
CTATTTACTTGAGTCTTGTACATACATACTTTTTTTATGAGCTATGAAATAAAACATTTTAAACTG
```

FIGURE 55A

```
CAATTCGGCCTCGCTCCTTGTGATTGCGCTAAACCTTCCGTCCTCAGCTGAGAACGCTCCACCACCTCCCCGGATCG
CTCATCTCTTGGCTGCCCTCCCACTGTTCCTGATGTTATTTTACTCCCCGTATCCCCTACTCGTTCTTCACAATTCT
GTAGGTGAGTGGTTCCAGCTGGTGCCTGGCCTGTGTCTCTTGGATGCCCTGTGGCTTCAGTCCGTCTCCTGTTGCCC
ACCACCTCGTCCCTGGGCCGCCTGATACCCCAGCCCAACAGCTAAGGTGTGGATGACAGTAGGGGGCTGGCTTCTC
TCACTGGTCAGGGGTCTTCTCCCCTGTCTGCCTCCCGGAGCTAGGACTGCAGAGGGGCCTATCATGGTGCTTGCAGG
CCCCCTGGCTGTCTCGCTGTTGCTGCCCAGCCTCACACTGCTGGTGTCCCACCTCTCCAGCTCCCAGGATGTCTCCA
GTGAGCCCAGCAGTGAGCAGCAGCTGTGCGCCCTTAGCAAGCACCCCACCGTGGCCTTTGAAGACCTGCAGCCGTGG
GTCTCTAACTTCACCTACCCTGGAGCCCGGGATTTCTCCCAGCTGGCTTTGGACCCCTCCGGGAACCAGCTCATCGT
GGGAGCCAGGAACTACCTCTTCAGACTCAGCCTTGCCAATGTCTCTCTTCTTCAGGCCACAGAGTGGGCCTCCAGTG
AGGACACGCGCCGCTCCTGCCAAAGCAAAGGGAAGACTGAGGAGGAGTGTCAGAAATACTGCGCGAGTCCTGATCGTC
GCCGGCCGGCAAGGTGTTCATGTGTGGAACCAATGCCTTTTCCCCCATGTGCACCAGCAGACAGGTGGGGAACCTCAG
CCGGACTATTGAGAAGATCAATGGTGTGGCCCGCTGCCCCTATGACCCCACGCCACAACTCCACAGCTGTCATCTCCT
CCCAGGGGGAGCTCTATGCAGCCACGGTCATCGACTTCTCCAGGTCGGGACCCTGCCATCTACCGCAGCCTGGGCAGT
GGGCCACCGCTTCGCACTGCCCCAATATAACTCCAAGTGGCTTAATGAGCCAAACTTCGTGGCAGCCTATGATATTGG
GCTGTTTGCATACTTCTTCCTGCGGGAGAACGCAGTGGAGCACGACTGTGGACGCACCGTGTACTCTCGCGTGGCCCC
GCGTGTGCAAGAATGACCGTGGGGGGCCCGATTCCTGCTGGAGGACACATGGACCACATTCATGAAGGCCCGGCTCAAC
TGCTCCCGCCCGGGCGAGGTCCCCTTCTACTATAACGAGCTGCAGAGTGCCTTCCACTTGCCGGGAGCAGGACCTCAT
CTATGGAGTTTTCACAACCAACGTAAACAGCATCGCCGGCTTCTGCTGTCTGCGCCTTCAACCTCAGTGCTATCTCCC
AGGCTTTCAATGGCCCATTTCGCTACCAGGAGAACCCCAGGGCTGCCTGGCTCCCCATAGCCAACCCCATCCCCAAT
TTCCAGTGTGGCACCCTGCCTGAGACCGGTCCCAACGAGAACCTGACGGAGCGCAGCCTGCAGGACGCGCAGCGCCT
CTTCCTGATGAGCGAGGCCGTGCAGCCGGTGACACCCGAGCCCTGTGTCACCCAGGACAGCGTGCGCTTCTCACACC
TCGTGGTGGACCTGGTGCAGGCTAAAGACACGCTCTACCATGTACTCTACATTGGCACCGAGTCGGGCACCATCCTG
AAGGCGCTGTCCACGGCGAGCCGCAGCCTCCACGGCTGCTACCTGGAGGAGCTGCACGTGCTGCCCCCCGGGCGCCG
CGAGCCCCTGCGCAGCCTGCGCATCCTGCACAGCGCCCGCGCGCTCTTCGTGGGGCTGAGAGACGGCGTCCTGCGGG
TCCCACTGGAGAGGTGCGCCGCCTACCGCAGCCAGGGGACGATGCCTGGGGGCCCGGGACCCGTACTGTGGCTGGGAC
GGGAAGCAGCAACGTTGCAGCACACTCGAGGACAGCTCCAACATGAGCCTCTGGACCCAGAACATCACCGCCTGTCC
TGTGCGGAATGTGACACGGGATGGGGGCTTCGGCCCATGGTCACCATGGCAACCATGTGAGCACTTGGATGGGGACA
ACTCAGGCTCTTGCCTGTGTCGAGCTCGATCCTGTGATTCCCCTCGACCCCGCTGTGGGGGCCTTGACTGCCTGGGG
CCAGCCATCCACATCGCCAACTGCTCCAGGAATGGGGCGTGGACCCCGTGGTCATCGTGGGCGCTGTGCAGCACGTC
CTGTGGCATCGGCTTCCAGGTCCGCCAGCGAAGTTGCAGCAACCCTGCTCCCCGCCACGGGGGCCGCATCTTCGTGG
GCAAGAGCCGGGAGGAACGGTTTCTGTAATGAGAACACGCCTTGCCCGGTGCCCATCTTCTCTGGGCTTCCTGGGGCTCC
TGGAGCAAGTGCAGCAGCAACTGTGGAGGGGGCATGCAGTCGCGGCGTCGGGCCTGCGAGAACGGCAACTCCTGCCT
GGGCTGCGGCGAGTTCAAGACGTGCAACCCCGAGGGCTGCCCCGAAGTGCGGCGCAACACCCCCTGGACGCCGTGGC
TGCCCGTGAACGTGACGCAGGGGGGGGCACGGCAGGAGCAGCGGTTCCGCTTCACCTGCCGCGCGCCCCTTGCGACAC
CCGCACGGCCTGCAGTTCGGCAGGAGAAGGACCGAGACGAGGACCTGTCCCGCGGACGGCTCCGGCTCCTGCGACAC
CGACGCCCTGGTGGAGGTCCTCCTGCGCAGCGGGAGCACCTCCCCGCACACGGTGAGCGGGGGCTGGGCCGCCTGGG
GCCCGTGGTCGTCCTGCTCCCGGGACTGCGAGCTGGGCTTCCGCCTCCGCAAGAGAACGTGCACTAACCCGCACCCCT
CGCAACGGGGGCCTGCCCTGCGTGGGCGATGCTGCCGAGTACCAGGACTGCAACCCCCAGGCTTGCCCAGTTCGGGG
TGCTTGGTCCTGCTGGACCTCATGGTCTCCATGCTCAGCTTCCTGTGGTGGGGGTCACTATCAACGCCACCCGTTCCT
GCACCAGCCCCGCACCCTCCCCAGGTGAGGACATCTGTCTCGGGCTGCACACGGAGGAGGCACTATGTGCCACACAG
GCCTGCCCAGGCTGGTCGCCCCTGGTCTGAGTGGAGTAAGTGCACTGACGACGGAGCCCAGAGCCGAAGCCGGCACTG
TGAGGAGCTCCTCCCAGGGTCCAGCGCCTGTGCTGGAAACAGCAGCCAGAGCCGCCCCTGCCCCTACAGCGAGATTC
CCGTCATCCTGCCAGCCTCCAGCACCCCACTCCAAAGTCTGGACTCTTTCCACATCCTGCTCCAGACAGCCAAGCTTTG
TTGGGGTCCCCACTGCTTTGAGATGGGTTCAATCTCATCCACTTGGTGGCCACGGGCATCTCCTGCTTCTTGGGCTC
TGGGCTCCTGACTGGACCTAGCCAGTGTACCTGTCTTGCCAGCACTGCCAGCGTCAGTCCCAGGAGTCCACACTGGTCCATC
CTGCCACCCCCAACCATTTGCACTACAAGGGCGGAGGCACCCCGAAGAATGAAAAGTACACACCCATGGAATTCAAG
ACCCTGAACAAGAATAACTTGATCCCTGATGACAGAGCCAACTTCTACCCATTGCAGCAGACCAATGTGTACACGAC
TACTTACTACCCAAGCCCCCTGAACAAACACAGCTTCCCGGCCCGAGGCCTCACCTGGACAACGGTGCTTCCCCAACA
GCTGATACCGCCGTCCTGGGGACTTGGGCTTCTTGCCTTCATAAGGCACAGAGCAGATGGAGATGGGACAGTGGAGAC
CAGTTTGGTTTTCTCCCTCTGCACTAGGCCAAGAACTTGCTGCCTTGCCTGTGGGGGGTCCCATCCGGCTTCAGAGA
GCTCTGGCTGGCATTGACCATGGGGGAAAGGGCTGGTTTCAGGCTGACATATGGCCGCAGGTCCAGTTC
```

FIGURE 55B

```
AGCCCAGGTCTCTCATGGTTATCTTCCAACCCACTGTCACGCTGACACTATGCTGCCATGCCTGGGCTGTGGACCTA
CTGGGCATTTGAGGAATTGGAGAATGGAGATGGCAAGAGGGCAGGCTTTTAAGTTTGGGTTGGAGACAACTTCCTGT
GGCCCCCACAAGCTGAGTCTGGCCTTCTCCAGCTGGCCCCAAAAAAGGCCTTTGCTACATCCTGATTATCTCTGAAA
GTAATCAATCAAGTGGCTCCAGTAGCTCTGGATTTTCTGCCAGGGCTGGGCCATTGTGGTGCTGCCCCAGTATGACA
TGGGACCAAGGCCAGCGCAGGTTATCCACCTCTGCCTGGAAGTCTATACTCTACCCAGGGCATCCCTCTGGTCAGAG
GCAGTGAGTACTGGGAACTGGAGGCTGACCTGTGCTTAGAAGTCCTTTAATCTGGGCTGGTACAGGCCTCAGCCTTG
CCCTCAATGCACGAAAGGTGGCCCAGGAGAGAGGATCAATGCCATAGGAGGCAGAAGTCTGGCCTCTGTGCCTCTAT
GGAGACTATCTTCCAGTTGCTGCTCAACAGAGTTGTTGGCTGAGACCTGCTTGGGAGTCTCTGCTGGCCCTTCATCT
GTTCAGGAACACACACACACACACACTCACACACGCACACACAATCACAATTTGCTACAGCAACAAAAAAGACATTG
GGCTGTGGCATTATTAATTAAAGATGATATCCAGTC
```

FIGURE 56

CGCAGAAAGAGGAGGCGCTTGCCTTCAGCTTGTGGGAAATCCCGAAG**ATG**GCCAAAGACAACTCAACTGTTCGTTGC
TTCCAGGGCCTGCTGATTTTTGGAAATGTGATTATTGGTTGTTGCGGCATTGCCCTGACTGCGGAGTGCATCTTCTT
TGTATCTGACCAACACAGCCTCTACCCACTGCTTGAAGCCACCGACAACGATGACATCTATGGGGCTGCCTGGATCG
GCATATTTGTGGGCATCTGCCTCTTCTGCCTGTCTGTTCTAGGCATTGTAGGCATCATGAAGTCCAGCAGGAAAATT
CTTCTGGCCGTATTTCATTCTGATGTTTATAGTATATGCCTTTGAAGTGGCATCTTGTATCACAGCAGCAACACAACA
AGACTTTTTCACACCCAACCTCTTCCTGAAGCAGATGCTAGAGAGGTACCAAAACAACAGCCCTCCAAACAATGATG
ACCAGTGGAAAAACAATGGAGTCACCAAAACCTGGGACAGGCTCATGCTCCAGGACAATTGCTGTGGCGTAAATGGT
CCATCAGACTGGCAAAAATACACATCTGCCTTCCGGACTGAGAATAATGATGCTGACTATCCCTGGCCTCGTCAATG
CTGTGTTATGAACAATCTTAAAGAACCTCTCAACCTGGAGGCTTGTAAACTAGGCGTGCCTGGTTTTTATCACAATC
AGGGCTGCTATGAACTGATCTCTGGTCCAATGAACCGACACGCCTGGGGGGTTGCCTGGTTTGGATTTGCCATTCTC
TGCTGGACTTTTTGGGTTCTCCTGGGTACCATGTTCTACTGGAGCAGAATTGAATAT**TAA**GCATAAAGTGTTGCCAC
CATACCTCCTTCCCCGAGTGACTCTGGATTTGGTGCTGGAACCAGCTCTCTCCTAATATTCCACGTTTGTGCCCCAC
ACTAACGTGTGTGTCTTACATTGCCAAGTCAGATGGTACGGACTTCCTTTAGGATCTCAGGCTTCTGCAGTTCTCAT
GACTCCTACTTTTCATCCTAGTCTAGCATTCTGCAACATTTATATAGACTGTTGAAAGGAGAATTTGAAAAATGCAT
AATAACTACTTCCATCCCTGCTTATTTTTAATTTGGGAAAATAAATACATTCGAAGGAACCTGTGTTATCACAGTAA
CCCAGAGCTGTATTTGGCTAGCAATCTGCCTGTATCTCTCACTATTATCTAAAAGAAACCTTCCAATGCTTCTGTTG
ATCTCAGTATTGTCAGGGGAACAGAGAAGTTGGGAAAAGATTACTGAAATATACCTTTTGCATTTCTTTCTAGAGTA
GCTCCCATATATGGAGATGGGTGATTCTCTTGATGCCACCTTCAGATCCTTTTATTCTCCAGAATAATTCTTAACAG
TGGTTCAAATTTCCTTTCATACCTTGAAGTATGTGTTTAGTAGCCTCAATTCTCCATTAATTAAAAGTGTGGGCTGG
GCGTGGGGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGGTCAGGAGTTCAAG
ACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACAAAAATACAAAAATTAGCCAGGCGTGATGGCAGGTGCCTGT
AATCCTAGCTACTTGGCAGGCTAACGCAGGAGAATCACTTGACCGGGAGACAGAGGTTGCAGTGAGCTGAGATCGTA
CCTATTGCACTCCATCCTGGATGAAAGAGCCAGACTCTGTCTCAAAACAAACAAAAAAGCGTGGGGACTTCTGGGGA
CAGACAAGGTGCCTGTTATATATTTACTCAGTCTTTGCCCTGAATGGTCTCAGCTTGAGACCATTTCAAACTGGAGA
GAAGCAAGCCAGCCAATAGAATGGGGTGATTTACAGGGATTTCTGTTTACTGTCAAAATATTTCTCATCTGCACTAT
GTTTCCATTTGTGGTCCTGAAGGAAATTCTTATAACTCAACATTTGTCTGGTCTTATAAGTAAAGACAGCTTTAAAA
TCTGTTCACTTTCAAA

FIGURE 57

```
CCCGGCCCGGGCTCGAGAATCAAGGGCCTCGGCCGCCGTCCCGCAGCTCAGTCCATCGCCCTTGCCGGGCAGCCCGG
GCAGAGACCATGTTTGACAAGACGCGGCTGCCGTACGTGGCCCTCGATGTGCTCTGCGTGTTGCTGGCTGGATTGCC
TTTTGCAATTTTTACTTCAAGGCATATTACTTCAAGGCATACCCCCTTCCAACGAGGAGTATTCTGTAATGATGAGT
CCATCAAGTACCCTTACAAAGAAGACACCATACCTTATGCGTTATTAGGTGGAATAATCATTCCATTCAGTATTATC
GTTATTATTCTTGGAGAAACCCTGTCTGTTTACTGTAACCTTTTGCACTCAAATTCCTTTATCAGGAATAACTACAT
AGCCACTATTTACAAAGCCATTGGAACCTTTTTATTTGGTGCAGCTGCTAGTCAGTCCCTGACTGACATTGCCAAGT
ATTCAATAGGCAGACTGCGGCCTCACTTCTTGGATGTTTGTGATCCAGATTGGTCAAAATCAACTGCAGCGATGGT
TACATTGAATACTACATATGTCGAGGGAATGCAGAAAGAGTTAAGGAAGGCAGGTTGTCCTTCTATTCAGGCCACTC
TTCGTTTTCCATGTACTGCATGCTGTTTGTGGCACTTTATCTTCAAGCCAGGATGAAGGGAGACTGGGCAAGACTCT
TACGCCCCACACTGCAATTTGGTCTTGTTGCCGTATCCATTTATGTGGGCCTTTCTCGAGTTTCTGATTATAAACAC
CACTGGAGCGATGTGTTGACTGGACTCATTCAGGGAGCTCTGGTTGCAATATTAGTTGCTGTATATGTATCGGATTT
CTTCAAAGAAAGAACTTCTTTTAAAGAAAGAAAAGAGGAGGACTCTCATACAACTCTGCATGAAACACCAACAACTG
GGAATCACTATCCGAGCAATCACCAGCCTTGAAAGGCAGCAGGGTGCCCAGGTGAAGCTGGCCTGTTTTCTAAAGGA
AAATGATTGCCACAAGGCAAGAGGATGCATCTTTCTTCCTGG
```

FIGURE 58

```
AGTGAAGGGGTTTCCCATATGAAAAATACAGAAAGAATTATTTGAATACTAGCAAATACACAACTTGATATTTCTAG
AGAACCCAGGCACAGTCTTGGAGACATTACTCCTGAGAGACTGCAGCTGATGGAAGATGAGCCCCAACTTCTAAAAA
TGTATCACTACCGGGATTGAGATACAAACAGCATTTAGGAAGGTCTCATCTGAGTAGCAGCTTCCTGCCCTCCTTCT
TGGAGATAAGTCGGGCTTTTGGTGAGACAGACTTTCCCAACCCTCTGCCCGGCCGGTGCCCATGCTTCTGTGGCTGC
TGCTGCTGATCCTGACTCCTGGAAGAGAACAATCAGGGGTGGCCCCAAAAGCTGTACTTCTCCTCAATCCTCCATGG
TCCACAGCCTTCAAAGGAGAAAAAGTGGCTCTCATATGCAGCAGCATATCACATTCCCTAGCCCAGGGAGACACATA
TTGGTATCACGATGAGAAGTTGTTGAAAATAAAACATGACAAGATCCAAATTACAGAGCCTGGAAATTACCAATGTA
AGACCCGAGGATCCTCCCTCAGTGATGCCGTGCATGTGGAATTTTCACCTGACTGGCTGATCCTGCAGGCTTTACAT
CCTGTCTTTGAAGGAGACAATGTCATTCTGAGATGTCAGGGGAAAGACAACAAAAACACTCATCAAAAGGTTTACTA
CAAGGATGGAAAACAGCTTCCTAATAGTTATAATTTAGAGAAGATCACAGTGAATTCAGTCTCCAGGGATAATAGCA
AATATCATTGTACTGCTTATAGGAAGTTTTACATACTTGACATTGAAGTAACTTCAAAACCCCTAAATATCCAAGTT
CAAGAGCTGTTTCTACATCCTGTGCTGAGAGCCAGCTCTTCCACGCCCATAGAGGGGAGTCCCATGACCCTGACCTG
TGAGACCCAGCTCTCTCCACAGAGGCCAGATGTCCAGCTGCAATTCTCCCTCTTCAGAGATAGCCAGACCCTCGGAT
TGGGCTGGAGCAGGTCCCCCAGACTCCAGATCCCTGCCATGTGGACTGAAGACTCAGGGTCTTACTGGTGTGAGGTG
GAGACAGTGACTCACAGCATCAAAAAAAGGAGCCTGAGATCTCAGATACGTGTACAGAGAGTCCCTGTGTCTAATGT
GAATCTAGAGATCCGGCCCACCGGAGGGCAGCTGATTGAAGGAGAAAATATGGTCCTTATTTGCTCAGTAGCCCAGG
GTTCAGGGACTGTCACATTCTCCTGGCACAAAGAAGGAAGAGTAAGAAGCCTGGGTAGAAAGACCCAGCGTTCCCTG
TTGGCAGAGCTGCATGTTCTCACCGTGAAGGAGAGTGATGCAGGGAGATACTACTGTGCAGCTGATAACGTTCACAG
CCCCATCCTCAGCACGTGGATTCGAGTCACCGTGAGAATTCCGGTATCTCACCCTGTCCTCACCTTCAGGGCTCCCA
GGGCCCACACTGTGGTGGGGGACCTGCTGGAGCTTCACTGTGAGTCCCTGAGAGGCTCTCCCCCGATCCTGTACCGA
TTTTATCATGAGGATGTCACCCTGGGGAACAGCTCAGCCCCCTCTGGAGGAGGAGCCTCCTTCAACCTCTCTCTGAC
TGCAGAACATTCTGGAAACTACTCCTGTGATGCAGACAATGGCCTGGGGGCCCAGCACAGTCATGGAGTGAGTCTCA
GGGTCACAGTTCCGGTGTCTCGCCCCGTCCTCACCCTCAGGGCTCCCGGGGCCCAGGCTGTGGTGGGGGACCTGCTG
GAGCTTCACTGTGAGTCCCTGAGAGGCTCCTTCCCGATCCTGTACTGGTTTTATCACGAGGATGACACCTTGGGGAA
CATCTCGGCCCACTCTGGAGGAGGGGCATCCTTCAACCTCTCTCTGACTACAGAACATTCTGGAAACTACTCATGTG
AGGCTGACAATGGCCTGGGGGCCCAGCACAGTAAAGTGGTGACACTCAATGTTACAGGAACTTCCAGGAACAGAACA
GGCCTTACCGCTGCGGGAATCACGGGGCTGGTGCTCAGCATCCTCGTCCTTGCTGCTGCTGCTGTGCTGCATTA
CGCCAGGGCCCGAAGGAAACCAGGAGGACTTTCTGCCACTGGAACATCTAGTCACAGTCCTAGTGAGTGTCAGGAGC
CTTCCTCGTCCAGGCCTTCCAGGATAGACCCTCAAGAGCCCACTCACTCTAAACCACTAGCCCCAATGGAGCTGGAG
CCAATGTACAGCAATGTAAATCCTGGAGATAGCAACCCGATTTATTCCCAGATCTGGAGCATCCAGCATACAAAAGA
AAACTCAGCTAATTGTCCAATGATGCATCAAGAGCATGAGGAACTTACAGTCCTCTATTCAGAACTGAAGAAGACAC
ACCCAGACGACTCTGCAGGGGAGGCTAGCAGCAGAGGCAGGGCCCATGAAGAAGATGATGAAGAAAACTATGAGAAT
GTACCACGTGTATTACTGGCCTCAGACCAC**TAG**CCCCTTACCCAGAGTGGCCCACAGGAAACAGCCTGCACCATTTT
TTTTTCTGTTCTCTCCAACCACACATCATCCATCTCTCCAGACTCTGCCTCCTACGAGGCTGGGCTGCAGGGTATGT
GAGGCTGAGCAAAAGGTCTGCAAATCTCCCCTGTGCCTGATCTGTGTGTTCCCCAGGAAGAGAGCAGGCAGCCTCTG
AGCAAGCACTGTGTTATTTTCACAGTGGAGACACGTGGCAAGGCAGGAGGGCCCTCAGCTCCTAGGGCTGTCGAATA
GAGGAGGAGGAGAAATGGTCTAGCCAGGGTTACAAGGGCACAATCATGACCATTTGATCCAAGTGTGATCGAAAGC
TGTTAATGTGCTCTCTGTATAAACAATTTGCTCCAAATATTTTGTTTCCCTTTTTTGTGTGGCTGGTAGTGGCATTG
CTGATGTTTTGGTGTATATGCTGTATCCTTGCTACCATATTGGG
```

FIGURE 59

```
GGAGCCGCGCCGCATCTCAGGCGCAGTCTCTAGGGGCTGTGCGCATCCTAGGGGGGGACATGTGCATCTCAGGGGGG
CTGCTCGCATCTGGGGGGTGCTGTGTGCATCTCGGGGGGGCTGTTGCATCTACGCGGGGTGGCTGTGTCCGCATCTG
GAGGGGGCTGTGCGCAACCCGGGGGGGGTGTTGCGCGCATCTAGCAGGGGCGGCTGTGCGCATTTCGGGGGGGGGCT
GTGCATATCTGGGGGGACACGTGCTTATCTCTGGGGGCGGCTGTGCGCATCTTGAGGGGTGTGTACATCTCGGGGGG
CCTGTGCATCTTGGGGGGCTGTGTGCATCCGCGGGGGCTGTGCGCATCTCGGGGTGCTGTGCGCTGCTCCTCTGAGC
TCTGCTCTTTCTTGCAGCGTTTGCCTCCAGCC**ATG**GAGGGCGGGGCCGCGGCAGCCACCCCCACAGCACTGCCTTACT
ACGTGGCCTTCTCCCAGCTGCTGGGCCTGACCTTGGTGGCCATGACCGGCGCGTGGCTCGGGCTGTACCGAGGCGGC
ATTGCCTGGGAGAGCGACCTGCAGTTCAACGCGCACCCCCTCTGCATGGTCATAGGCCTGATCTTCCTGCAGGGAAA
TGCCCTGCTGGTTTACCGTGTCTTCAGGAACGAAGCTAAACGCACCACCAAGGTCCTGCACGGGCTGCTGCACATCT
TTGCGCTCGTCATCGCCCTGGTTGGCTTGGTGGCGGTGTTCGACTACCACAGGAAGAAGGGCTACGCTGACCTGTAC
AGCCTACACAGCTGGTGCGGGATCCTTGTCTTTGTCCTGTACTTTGTGCAGTGGCTGGTGGGCTTCAGCTTCTTCCT
GTTCCCCGGAGCTTCATTCTCCCTGCGGAGCCGCTACCGCCCACAGCACATCTTCTTTGGTGCTACCATCTTCCTCC
TTCCCGTGGGCACCGCCCTGCCTGGGCCTGAAGGAGGCACTGCTGTTCAACCTCGGGGGCAAGTATAGCGCATTTGAG
CCCGAGGGTGTCCTGGCCAACGTGCTGGGCCTGCTGCTGGCCTGCTTCGGTGGGGCGGTGCTCTACATCTTGACCCG
GGCCGACTGGAAGCGGCCTTCCCAGGCGGAAGAGCAGGCCCTCTCCATGGACTTCAAGACGCTGAGGCAGGGAGATA
GCCCCGGCTCCCAG**TGA**TGCGCCCGGCCCGGCCCTGGGGGTTCGCGGGGTGTCTTCTTGCCTGCCCCTGCTGAGGCGT
CTTCAGGACTGCAGGCTCCGGAGAGTGGCTCTGGCAGCAGGCGGGCGCGTGGGTGCAGGGGGATCCGTTTGATGCGT
CGTTTCTGGGGCAGGTCTCCGCCTCCTCTGCTTCTCGTTTCTCCGCTGCTATAGACCAGTTCATTGTGTGTGGCTCC
CGTGTCTCTGTTGCCCCCTTCAGTGCAGAAGGCTTTGGGTAGACTTCGGGTGTTCGGTCCTGGTCGCAGAGCACAGA
TCTTTAAAGAAGCGTTAGAGAGGTAGGTTCTACCCTCTTGGTAGTAGATGCCTGGGGCAAGGCCCAGGGGAAACTGG
GGGGGCCTCAGGGACAGGCCTGGAAAGGCCACGATGGCCTGCTGAATTCAAACAAGGAGTCCCTCCAGCCTGAATAA
CACGTGGCACAAATGGGCCCGGCCTTTGGCAGAGGAGCAAGTGATATGATGTGTAAAGTATGTTGGTGGTGAAAGCA
AGGTTCCCCAGGAGAGGGGAGGGACTGGCCCCTGGGAAGCTGTGAGATGAGGCTGTGGCCCAGCTGTAGTCCTGACC
TTACTCTTCTTTAAAACCCTTTAGCCCTAGGATGGCTTTGGTGGGAGAGGGGATAGAAGCCCATGACTTCAGACAGA
CTTTCTCTTGGCAGATGCAGGCAGGCCTCCTCCCAGGCTGCTCCAGACATGGGGGTTGGGGATGGGGGGTACCTTGC
AGCCCCTTCCTGCTGGGGCTCCCTCCTTGTAGCACCCCCTTGCGGCTCAGCTCTGGTTTCCTCTCCCAGGCTCACCC
AGGCTCTGCTCAGGCTGGGAGGCAGAGGGCACAAACCTTATAATTTTTAAATGAAAAACCGCTGCTGCTGGCTGTG
GCTAGAGCCCCCTGGGGCTGCTGGAGCTGCTGCCTCTGTTCTGGAGGACGAGCCTTCTCCTTATCTGCTGCCCATCT
TTCCAGGAAGTCAGGATGGAGTCAGACAACTAACGATCATCCCCCGTGGGTGTCTGCACATCACTCCAGCCCCATAA
AGAGTGTCATGTTAGCTGAGTCACCATTTGGCTTCGGCCTGGAAATAGTGTGTTAGAACACTGATCGTGTGCGAGGC
CAGGAGATCAAGACCATCCTGACTAACAAACACAGTGAAACCCCGTCTCTACTAAAAATACAAAAAAATTAGCCAGG
CGTGGTGGTGGGCGCCTGTAGTCCCAGCTACTTGGGAGACAGGTCTCTACGGAATTCCCTGTATTAGTCTATATGGT
TCTCCAAGAAACTGAATGAATCCATTGGAGAAGCGGTGGATAACTAGCCACGACAAAATTTGAGAATACATAAACAA
CGCATTCCGCAGGAAACATACAGAGGATGCCTTTTCTGTGATTGGGTGGGATTTTTTCCCTTTTTATGTGGATATAG
TAGTTACTTGTGACAAGAATAATTTTGGAATAATTTCTATTAATATCAACTCTGAAGCTAATTGTACATAATCTCGA
GATTGTGTTTGTTCATAATAAAAGTGAAGTGAATGTG
```

FIGURE 60

CGCCACCGCTGGGTGCGGCGAGGCCGGCGCG**ATG**CGGCAGCTGTGCCGGGGCCGCGTGCTGGGCATCTCGGTGGCCA
TCGCGCACGGGGTCTTCTCGGGCTCCCTCAACATCTTGCTCAAGTTCCTCATCAGCCGCTACCAGTTCTCCTTCCTG
ACCCTGGTGCAGTGCCTGACCAGCTCCACCGCGGCGCTGAGCCTGGAGCTGCTGCGGCGCCTCGGGCTCATCGCCGT
GCCCCCCTTCGGTCTGAGCCTGGCGCGCTCCTTCGCGGGGGTCGCGGTGCTCTCCACGCTGCAGTCCAGCCTCACGC
TCTGGTCCCTGCGCGGCCTCAGCCTGCCCATGTACGTGGTCTTCAAGCGCTGCCTGCCCCTGGTCACCATGCTCATC
GGCGTCCTGGTGCTCAAGAACGGCGCGCCCTCGCCAGGGGTGCTGGCGGCGGTGCTCATCACCACCTGCGGCGCCGC
CCTGGCAGGAGCCGGCGACCTGACGGGCGACCCCATCGGGTACGTCACGGGAGTGCTGGCGGTGCTGGTGCACGCTG
CCTACCTGGTGCTCATCCAGAAGGCCAGCGCAGACACCGAGCACGGGCCGCTCACCGCGCAGTACGTCATCGCCGTC
TCTGCCACCCCGCTGCTGGTCATCGCTCCTTCGCCAGCACCGACTCCATCCACGCCTGGACCTTCCCGGGCTGGAA
GGACCCGGCCATGGTCTGCATCTTCGTGGCCTGCATCCTGATCGGCTGCGCCATGAACTTCACCACGCTGCACTGCA
CCTACATCAATTCGGCCGTGACCACCTCTCTGTTCATTGCCGGCGTGGTGGTGAACACCCCTGGGCTCTATCATTTAC
TGTGTGGCCAAGTTCATGGAGACCAGAAAGCAAAGCAACTACGAGGACCTGGAGGCCCAGCCTCGGGGAGAGGAGGC
GCAGCTAAGTGGAGACCAGCTGCCGTTCGTGATGGAGGAGCTGCCCGGGGAGGGAGGAAATGGCCGGTCAGAAGGTG
GGGAGGCAGCAGGTGGCCCCGCTCAGGAGAGCAGGCAAGAGGTCAGGGGCAGCCCCCGAGGAGTCCCGCTGGTGGCT
GGGAGCTCTGAAGAAGGGAGCAGGAGGTCGTTAAAAGATGCTTACCTCGAGGTATGGAGGTTGGTTAGGGGAACCAG
GTATATGAAGAAGGATTATTTGATAGAAAACGAGGAGTTACCCAGTCCT**TGA**GAAGGAGGTGCATGTACGTACCTAT
GTGCATACACTTATTTTATATGTTAGAAATGACGTGTTTTAATGAGAGGCCTCCCCGTTTTATTCTTTGAGGAGTGG
GGAAGGGAAGAAAAGAAAGAAGCTGAAAGGTACTGACACAGAGCAACAAAATTAGCACCTGTGTGAATTATTTAGTG
TGACTTCACCTGAGGCATCACAGAGACAAAAGAATGTGAAGCTACTTAACAAAGTAAGGCAACGTTTCTGCTTCAGA
CTCCTGGCACATTTACTTTTTGTCATTATAACCATAACTAAATATCTGCATGTACCAAGAGTCCCTAAGCCACCCCC
TCCAAAGATGGAGTGTAGAAATGATGACAGCACTTAGTAAGTTCAAAGATGACATTCAGGGATGCATTTTTTGATGA
TAGAACTACAGTTTTTATCGCCAGCTGGGCAAAGAGTATATTGCTGAAATGATATATAAATATATTGAATTGATGTT
TACTGTTTATAGTCATCTGAAATATCATATTTACTCTGATTCTACTCACTTGTTTTTTAAAAATAAGTGTCCTACTA
TTGTATTATATATTGATAGAAACTGTTAAAGCTATTTTGAAAATATGAGTTCTTAGCTTTAATCATGAAGTCTGAAG
TTTGCTTTCAGTAATTATTTTAAAAGTTGTTTTGGTTCATTGCTTTATAATATTTATTATTGAATGCCAAACCTGTT
CTTTTTTTTACTGTGTCCAATATTCTTTCAAGCAAATGCAATGGCTGGAATATAATTCAGAATTAACTGAAACCCAG
CCAGAAGAGGGACCACCTGTAAAGCAAGTCCTTTCAAGTTTCACTGCACATCCCAAACCATGTTACAAAAAGAGCAA
CTGCTATATTCACATTATGATATTTTTCTATCTTAAATTTGTCAAAATAAAGTATGAGTCTAACTATTAAAAAAAAA
AAAAAAAAA

FIGURE 61

```
ATTTCTCCAGGAGAACTCCACACTTGCACAATATCTTCTCCAAGGACAGAAGAGGGACAATATTTGTCCTTCCTACT
GGATTTTGAAACCCTTTGCACTGAGTGTAGACTGTAGTTCTGTGACATACCATGAAGAGTGTGTGTGTGTTTTAAGG
GAGCTACTGTCTTACCCAAAACCTGTGAATATAAAGTGTTTTTTCATGAATTGCTCATTATTCAGCCAGTCGTTAAT
GAATTCATTCAACAAGTGTCTCTGAGATGCTAGACACTGGGGATTAAAAGAGGAACAACAGAGACAAGATCTCTGCC
CTCCAGAAACTGACAGTCTATTGAATGAGACAGTTGTCTAACAATCACAATCAAGTGTGATCAATCTCTGGTGACAG
GACTCTAACCTAGAGGCATGTGCCTAATCTGGGGTGACT
```

FIGURE 62

CAGGGGACAGGCTGCAGCCGGTGCAGTTACACGTTTTCCTCCAAGGAGCCTCGGACGTTGTCACGGGTTTGGGGTCG
GGGACAGAGCAGTGACC**ATG**GCCAGGCTGGCGTTGTCTCCTGTGCCCAGCCACTGGATGGTGGCGTTGCTGCTGCTG
CTCTCAGCTGAGCCAGTACCAGCAGCCAGATCGGAGGACCGGTACCGGAATCCCAAAGGTAGTGCTTGTTCGCGGAT
CTGGCAGAGCCCACGTTTCATAGCCAGGAAACGGGGCTTCACGGTGAAAATGCACTGCTACATGAACAGCGCCTCCG
GCAATGTGAGCTGGCTCTGGAAGCAGGAGATGGACGAGAATCCCCAGCAGCTGAAGCTGGAAAAGGGCCGCATGGAA
GAGTCCCAGAACGAATCTCTCGCCACCCTCACCATCCAAGGCATCCGGTTTGAGGACAATGGCATCTACTTCTGTCA
GCAGAAGTGCAACAACACCTCGGAGGTCTACCAGGGCTGCGGCACAGAGCTGCGAGTCATGGGATTCAGCACCTTGG
CACAGCTGAAGCAGAGGAACACGCTGAAGGATGGTATCATCATGATCCAGACGCTGCTGATCATCCTCTTCATCATC
GTGCCTATCTTCCTGCTGCTGGACAAGGATGACAGCAAGGCTGGCATGGAGGAAGATCACACCTACGAGGGCCTGGA
CATTGACCAGACAGCCACCTATGAGGACATAGTGACGCTGCGGACAGGGGAAGTGAAGTGGTCTGTAGGTGAGCACC
CAGGCCAGGAG**TGA**GAGCCAGGTCGCCCCATGACCTGGGTGCAGGCTCCCTGGCCTCAGTGACTGCTTCGGAGCTGC
CTGGCTCATGGCCCAACCCCTTTCCTGGACCCCCCAGCTGGCCTCTGAAGCTGGCCCACCAGAGCTGCCATTTGTCT
CCAGCCCCTGGTCCCCAGCTCTTGCCAAAGGGCCTGGAGTAGAAGGACAACAGGGCAGCAACTTGGAGGGGAGTTCTC
TGGGGATGGACGGGACCCAGCCTTCTGGGGGTGCTATGAGGTGATCCGTCCCCACACATGGGATGGGGGGAGGCAGAG
ACTGGTCCAGAGCCCGCAAATGGACTCGGAGCCGAGGGCCTCCCAGCAGAGCTTGGGAAGGGCCATGGACCCAACTG
GGCCCCAGAAGAGCCACAGGAACATCATTCCTCTCCCGCAACCACTCCCACCCCAGGGAGGCCCTGGCCTCCAGTGC
CTTCCCCCCGTGGAATAAACGGTGTGTCCTGAGAAACCA

FIGURE 63

ATGGGCCTCCCCGAGCCGGGCCCTCTCCGGCTTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCT
CCAGCATCTTGCGGCGGCAGCGGCTGATCCGCTGCTCGGCGGCCAAGGGCCGGCCAAGGAGTGCGAAAAGGACCAAT
TCCAGTGCCGGAACGAGCGCTGCATCCCCTCTGTGTGGAGATGCGACGAGGACGATGACTGCTTAGACCACACAGCGAC
GAGGACGACTGCCCCAAGAAGACCTGTGCAGACAGTGACTTCACCTGTGACAACGGCCACTGCATCCACGAACGGTG
GAAGTGTGACGGCGAGGAGGAGTGTCCTGATGGCTCCGATGAGTCCGAGGCCACTTGCACCAAGCAGGTGTGTCCTG
CAGAGAAGCTGAGCTGTGGACCCACCAGCCACAAGTGTGTACCTGCCTCGTGGCGCTGCGACGGGGAGAAGGACTGC
GAGGGTGGAGCGGATGAGGCCGGCTGTGCTACCTCACTGGGCACCTGCCGTGGGGACGAGTTCCAGTGTGGGGATGG
GACATGTGTCCTTGCAATCAAGCACTGCAACCAGGAGCAGGACTGTCCAGATGGGAGTGATGAAGCTGGCTGCCTAC
AGGGGCTGAACGAGTGTCTGCACAACAATGGCGGCTGCTCACACATCTGCACTGACCTCAAGATTGGCTTTGAATGC
ACGTGCCCAGCAGGCTTCCAGCTCCTGGACCAGAAGACTTGTGGCGACATTGATGAGTGCAAGGACCCAGATGCCTG
CAGCCAGATCTGTGTCAATTACAAGGGCTATTTTAAGTGTGAGTGCTACCCTGGCTGCGAGATGGACCTACTGACCA
AGAACTGCAAGGCTGCTGCTGGCAAGAGCCCATCCCTAATCTTCACCAACCGCACGAGTGCGGAGGATCGACCTGTG
AAGCGGAACTATTCACGCCTCATCCCCATGCTCAAGAATGTCGTGGCACTAGATGTGGAAGTTGCCACCAATCGCAT
CTACTGGTGTGACCTCTCCTACCGTAAGATCTATAGCGCCTACATGGACAAGGCCAGTGACCCGAAAGAGCGGGAGG
TCCTCATTGACGAGCAGTTGCACTCTCCAGAGGGCCTGGCAGTGGACTGGGTCCACAAGCACATCTACTGGACTGAC
TCGGGCAATAAGACCATCTCAGTGGCCACAGTTGATGGTGGCCGCCGACGCACTCTCTTCAGCCGTAACCTCAGTGA
ACCCCGGGCCATCGCTGTTGACCCCCTGCGAGGGTTCATGTATTGGTCTGACTGGGGGGACCAGGCCAAGATTGAGA
AATCTGGGCTCAACGGTGTGGACCGGCAAACACTGGTGTCAGACAATATTGAATGGCCCAACGGAATCACCCTGGAT
CTGCTGAGCCAGCGCTTGTACTGGGTAGACTCCAAGCTACACCAACTGTCCAGCATTGACTTCAGTGGAGGCAACAG
AAAGACGCTGATCTCCTCCACTGACTTCCTGAGCCACCCTTTTGGGATAGCTGTGTTTGAGGACAAGGTGTTCTGGA
CAGACCTGGAGAACGAGGCCATTTTCAGTGCAAATCGGCTCAATGGCCTGGAAATCTCCATCCTGGCTGAGAACCTC
AACAACCCACATGACATTGTCATCTTCCATGAGCTGAAGCAGCCAAGAGCTCCAGATGCCTGTGAGCTGAGTGTCCA
GCCTAATGGAGGCTGTGAATACCCTGTGCCTTCCTGCTCCTCAGATCTCCAGCCACTCTCCCAAGTACACATGTGCCT
GTCCTGACACAATGTGGCTGGGTCCAGCATGAAGAGGTGCTACCGAGATGCAAATGAAGACAGTAAGATGGGCTCA
ACAGTCACTGCCGCTGTTATCGGGATCATCGTGCCCCATAGTGGTGATAGCCCTCCTGTGCATGAGTGGATACCTGAT
CTGGAGAAACTGGAAGCGGAAGAACACCAAAAGCATGAATTTTGACAACCCAGTCTACAGGAAAACAACAGAAGAAG
AAGATGAAGATGAGCTCCATATAGGGAGAACTGCTCAGATTGGCCATGTCTATCCTGCACGAGTGGCATTAAGCCTT
GAAGATGATGGACTACCC**TGA**GGATGGGATCACCCCCTTCGTGCCTCATGGAATTCAGTCCCATGCACTACACTCCG
GATGGTGTATGACTGGATGAATGGGTTTCTATATATGGGTCTGTGTGAGTGTATGTGTGTGTGTGATTTTTTTTTTT
AAATTTATGTTGCGGAAAGGTAACCACAAAGTTATGATGAACTGCAAACATCCAAAGGATGTGAGAGTTTTTCTATG
TATAATGTTTTATACACTTTTTAACTGGTTGCACTACCCATGAGGAATTCGTGGAATGGCTACTGCTGACTAACATG
ATGCACATAACCAAATGGGGGCCAATGGCACAGTACCTTACTCATCATTTAAAAACTATATTTACAGAAGATGTTTG
GTTGCTGGGGGGCTTTTTTAGGTTTTGGGCATTTGTTTTTTGTAAATAAGATGATTATGCTTTGTGGCTATCCATCA
ACATAAGT

FIGURE 64

CCGTTCCGCGCTCTGGCGGCTCCTCCCGGGCG**ATG**CCTCCGCTCTGGGCCCTGCTGGCCCTCGGCTGCCTGCGGTTC
GGCTCGGCTGTGAACCTGCAGCCCCAACTGGCCAGTGTGACTTTCGCCACCAACAACCCCACACTTACCACTGTGGC
CTTGGAAAAGCCTCTCTGCATGTTTGACAGCAAAGAGGCCCTCACTGGCACCCACGAGGTCTACCTGTATGTCCTGG
TCGACTCAGCCATTTCCAGGAATGCCTCAGTGCAAGACAGCACCAACACCCCACTGGGCTCAACGTTCCTACAAACA
GAGGGTGGGAGGACAGGTCCCTACAAAGCTGTGGCCTTTGACCTGATCCCCTGCAGTGACCTGCCCAGCCTGGATGC
CATTGGGGATGTGTCCAAGGCCTCACAGATCCTGAATGCCTACCTGGTCAGGGTGGGTGCCAACGGGACCTGCCTGT
GGGATCCCAACTTCCAGGGCCTCTGTAACGCACCCCTGTCGGCAGCCACGGAGTACAGGTTCAAGTATGTCCTGGTC
AATATGTCCACGGGCTTGGTAGAGGACCAGACCCTGTGGTCGGACCCCATCCGCACCAACCAGCTCACCCCCATACTC
GACGATCGACACGTGGCCAGGCCGGCGGAGCGGAGGCATGATCGTCATCACTTCCATCCTGGGCTCCCTGCCCTTCT
TTCTACTTGTGGGTTTTGCTGGCGCCATTGCCCTCAGCCTCGTGGACATGGGGAGTTCTGATGGGGAAACGACTCAC
GACTCCCAAATCACTCAGGAGGCTGTTCCCAAGTCGCTGGGGGCCTCGGAGTCTTCCTACACGTCCGTGAACCGGGG
GCCGCCACTGGACAGGGCTGAGGTGTATTCCAGCAAGCTCCAAGAC**TGA**GCCCAGCACCACCCCTGGGCAGCAGCAT
CCTCCTCTCTGGCCTTGCCCCAGGCCCTGCAGCCGGTGGTTGTCACACCCTGACTTCAGGGAAGGTGAAACAGGGCTT
GTCCCTCCAACTGCAGGAAAACCCTTAATAAAATCTTCTGATGAGTTCTAAAAAAAAAA

FIGURE 65

TCTCTGACCACCGGTGCATGCAGCCCCTGTCACATACCGCCTGCTTGCTCAAATCAATCATGACCCTTTCATGTGAA
ATCTTTAGTATTGTGAGCCCTTAAAAGGGACGGAAATTGTGCATACGTGGAGCTCGGATTTTAAGGCAGTAGCTGCC
CGATGCTCCCAGCTGAATAAAGCCCTTCCTTCTACAATTTGGTGTCTGAGGGGTTTTGTCTGCGGCTCGTCCTGCTA
CATTTCTTGGTTCCCTGACCGGGAAGCAAGGTGACTGACAGACGGCCGAGGCAGCCCCTTAGGCGGCTTAAGCCTGC
CCTGTGGAGCATCCCTGCGGGGGACTCCGGCCAGCCTGAGTGACGCGATCCAAAGAGTGCTCCCGGGTAGGAAATTG
CCCCGGTGGAACGCCTCGCCAGAGCAGCGTGTAGCAGGCCCCCGCGGAGGATTAACACAGTGGCTGAACACCGGGAA
GGAACTGGCACTTGGAGTCCAGACATCTGAAACTTGACTGGGAGCTGTACGTGGATGGGAGCAGCTTCACCAACCCC
TGCAAAGTGACTCTGAAGAAGACGACAAGCCCTGCTCCAGTCACACCCGGAAGCTGACTGGTCCACGCACGGCCGAA
GC**ATG**AGGAAACTCATCGCAGGACTCATTTTTCTTAAATTTTGGACTTATACAGTAAGGGCTTCAACTGACCTTCCT
CAGACTGAGAACTGTTTCCAGTATATACATCAAGTCACTGAGATCTCCAGCACCCTGCCGGTGGCACTACTGAGAGA
CGAGGTGCCAGGGTGGTTCCTGAAAGTGCCTGAGCCCCAACTTATCAGCAAGGAGCTCATCATGCTGACAGAAGTCA
TGGAGGTCTGGCATGGCTTAGTGATCGCGGTGGTGTCCCTCTTCCTGCAGGCCTGCTTCCTCACCGCCATCAACTAC
CTGCTCAGCAGGCACATGGCCCACAAGAGTGAACAGATACTGAAAGCGGCCAGTCTCCAGGTTCCCAGGCCCAGCCC
TGGCCACCATCATCCACCTGCTGTCAAAGAGATGAAGGAGACTCAGACAGAGAGAGACATCCCAATGTCTGATTCCC
TTTACAGGCATGACAGCGACACACCCTCAGATAGCTTGGATAGCTCCTGCAGTTCGCCTCCTGCCTGCCAGGCCACA
GAGGATGTGGATTACACACAAGTCGTCTTTTCTGACCCTGGAGAACTAAAA**TGA**TCCCCGCTGGACTATGAGAACA
TAAAGGAAATCACAGATTATGTCAATGTCAATCCAGAAAGACACAAGCCCAGTTTCTGGTATTTTGTCAACCCTGCT
TCTTGTCTGAGCCAGCGGAATTATGATCAAGTGGCCATGTGAATTCCAAATATTTTTTAAATGGGGTCCAGTTCTCT
ATGGNTTTCTTANAATTTAATTTTGTAGGGGNAANTGCCATTTTNCCCCTTTTAAACAANGGNTTGGGGNTAAAAGN
TTTTTNGGGCCA

FIGURE 66

AAACTTTTGCAGCGGCTGGGTAGCAGCACGTCTCTTGCTCCTCAGGGGCCACTGCCAGGCTTGCCGAGTCCTGGGACT
GCTCTCGCTCCGGCTGCCACTCTCCCGCGCTCTCCTAGCTCCCTGCGAAGCAGG**ATG**GCCGGGACCGTGCGCCACCGC
GTGCTTGGTGGTGGCGATGCTGCTCAGCTTGGACTTCCCGGGACAGGCGCAGCCCCCGCCGCCGCCGCCGGACGGCCA
CCTGTCACCAAGTCCGCTCCTTCTTCCAGAGACTGCAGCCCGGACTCAAGTGGGTGCCAGAAACTCCCGTGCCAGGA
TCAGATTTGCAAGTATGTCTCCCTAAGGGCCCAACATGCTGCTCAAGAAAGATGGAAGAAAAATACCAACTAACAGC
ACGATTGAACATGGAACAGCTGCTTCAGTCTGCAAGTATGGAGCTCAAGTTCTTAATTATTCAGAATGCTGCGGTTT
TCCAAGAGGCCTTTGAAATTGTTGTTCGCCATGCCAAGAACTACACCAATGCCATGTTCAAGAACAACTACCCAAGC
CTGACTCCACAAGCTTTTGAGTTTGTGGGTGAATTTTTCACAGATGTGTCTCTCTACATCTTGGGTTCTGACATCAA
TGTAGATGACATGGTCAATGAATTGTTTGACAGCCTGTTTCCAGTCATCTATACCCAGCTAATGAACCCAGGCCTGC
CTGATTCAGCCTTGGACATCAATGAGTGCCTCCGAGGAGCAAGACGTGACCTGAAAGTATTTGGGAATTTCCCCAAG
CTTATTATGACCCAGGTTTCCAAGTCACTGCAAGTCACTAGGATCTTCCTTCAGGCTCTGAATCTTGGAATTGAAGT
GATCAACACAACTGATCACCTGAAGTTCAGTAAGGACTGTGGCCGAATGCTCACCAGAATGTGGTACTGCTCTTACT
GCCAGGGACTGATGATGGTTAAACCCTGTGGCGGTTACTGCAATGTGGTCATGCAAGGCTGTATGGCAGGTGTGGTG
GAGATTGACAAGTACTGGAGAGAATACATTCTGTCCCTTGAAGAACTTGTGAATGGCATGTACAGAATCTATGACAT
GGAGAACGTACTGCTTGGTCTCTTTTCAACAATCCATGATTCTATCCAGTATGTCCAGAAGAATGCAGGAAAGCTGA
CCACCACTATTGGCAAGTTATGTGCCCATTCTCAACAACGCCAATATAGATCTGCTTATTATCCTGAAGATCTCTTT
ATTGACAAGAAAGTATTAAAAGTTGCTCATGTAGAACATGAAGAAACCTTATCCAGCCGAAGAAGGGAACTAATTCA
GAAGTTGAAGTCTTTCATCAGCTTCTATAGTGCTTTGCCTGGCTACATCTGCAGCCATAGCCCTGTGGCGGAAAACG
ACACCCTTTGCTGGAATGGACAAGAACTCATGGAGAGATACAGCCAAAAGGCAGCAAGGAATGGAATGAAAAACCAG
TTCAATCTCCATGAGCTGAAAATGAAGGGCCCTGAGCCAGTGGTCAGTCAAATTATTGACAAACTGAAGCACATTAA
CCAGCTCCTGAGAACCATGTCTATGCCCAAAGGTAGAGTTCTGGATAAAAACCTGGATGAGGAAGGGTTTGAAAGTG
GAGACTGCGGTGATGATGAAGATGAGTGCATTGGAGGCTCTGGTGATGGAATGATAAAAGTGAAGAATCAGCTCCGC
TTCCTTGCAGAACTGGCCTATGATCTGGATGTGGATGATGCGCCTGGAAACAGTCAGCAGGCAACTCCGAAGGACAA
CGAGATAAGCACCTTTCACAACCTCGGGAACGTTCATTCCCCGCTGAAGCTTCTCACCAGCATGGCCATCTCGGTGG
TGTGCTTCTTCTTCCTGGTGCAC**TGA**CTGCCTGGTGCCCAGCACATGTGCTGCCCTACAGCACCCTGTGGTCTTCCT
CGATAAAGGGAACCACTTTCTTATTTTTTTCTATTTTTTTTTTTTTGTTATCCTGTATACCTCCTCCAGCCATGAAG
TAGAGGACTAACCATGTGTTATGTTTTCGAAAATCAAATGGTATCTTTTGGAGGAAGATACATTTTAGTGGTAGCAT
ATAGATTGTCCTTTTGCAAAGAAAGAAAAAAAAACCATCAAGTTGTGCCAAATTATTCTCCTATGTTTGGCTGCTAGA
ACATGGTTACCATGTCTTTCTCTCTCACTCCCTCCCTTTCTATCGTTCTCTCTTTGCATGGATTTCTTTGAAAAAAA
ATAAATTGCTCAAATA

FIGURE 67

```
GCGGAACACCGGCCCGCCGTCGCGGCAGCTGCTTCACCCCTCTCTCTGCAGCCATGGGGCTCCCTCGTGGACCTCTC
GCGTCTCTCCTCCTTCTCCAGGTTTGCTGGCTGCAGTGCGCGGCCTCCGAGCCGTGCCGGGCGGTCTTCAGGGAGGC
TGAAGTGACCTTGGAGGCGGGAGGCGCGGAGCAGGAGCCCGGCCAGGCGCTGGGGAAAGTATTCATGGGCTGCCCTG
GGCAAGAGCCAGCTCTGTTTAGCACTGATAATGATGACTTCACTGTGCGGAATGGCGAGACAGTCCAGGAAAGAAGG
TCACTGAAGGAAAGGAATCCATTGAAGATCTTCCCATCCAAACGTATCTTACGAAGACACAAGAGAGATTGGGTGGT
TGCTCCAATATCTGTCCCTGAAAATGGCAAGGGTCCCTTCCCCCAGAGACTGAATCAGCTCAAGTCTAATAAAGATA
GAGACACCAAGATTTTCTACAGCATCACGGGGCCGGGGGCAGACAGCCCCCCTGAGGGTGTCTTCGCTGTAGAGAAG
GAGACAGGCTGGTTGTTGTTGTTGAATAAGCCACTGGACCGGGAGGAGATTGCCAAGTATGAGCTCTTTGGCCACGCTGT
GTCAGAGAATGGTGCCTCAGTGGAGGACCCCATGAACATCTCCATCATCGTGACCGACCAGAATGACCACAAGCCCA
AGTTTACCCAGGACACCTTCCGAGGGAGTGTCTTAGAGGGAGTCCTACCAGGTACTTCTGTGATGCAGGTGACAGCC
ACAGATGAGGATGATGCCATCTACACCTACAATGGGGTGGTTGCTTACTCCATCCATAGCCAAGAACCAAAGGACCC
ACACGACCTCATGTTCACAATTCACCGGAGCACAGGCACCATCAGCGTCATCTCCAGTGGCCTGGACCGGGAAAAAG
TCCCTGAGTACACACTGACCATCCAGGCCACAGACATGGATGGGGACGGCTCCACCACCACGGCAGTGGCAGTAGTG
GAGATCCTTGATGCCAATGACAATGCTCCCATGTTTGACCCCCAGAAGTACGAGGCCCATGTGCCTGAGAATGCAGT
GGGCCATGAGGTGCAGAGGCTGACGGTCACTGATCTGGACGCCCCCAACTCACCAGCGTGGCGTGCCACCTACCTTA
TCATGGGCGGTGACGACGGGGACCATTTTACCATCACCACCCACCCTGAGAGCAACCAGGGCATCCTGACAACCAGG
AAGGGTTTGGATTTTGAGGCCAAAAACCAGCACACCCTGTACGTTGAAGTGACCAACGAGGCCCCTTTTGTGCTGAA
GCTCCCAACCTCCACAGCCACCATAGTGGTCCACGTGGAGGATGTGAATGAGGCACCTGTGTTTGTCCCACCCTCCA
AAGTCGTTGAGGTCCAGGAGGGCATCCCCACTGGGGAGCCTGTGTGTGTGTCTACACTGCAGAAGACCCTGACAAGGAG
AATCAAAAGATCAGCTACCGCATCCTGAGAGACCCAGCAGGGTGGCTAGCCATGGACCCAGACAGTGGGCAGGTCAC
AGCTGTGGGCACCCTCGACCGTGAGGATGAGCAGTTTGTGAGGAACAACATCTATGAAGTCATGGTCTTGGCCATGG
ACAATGGAAGCCCTCCCACCACTGGCACGGGAACCCTTCTGCTAACACTGATTGATGTCAACGACCATGGCCCAGTC
CCTGAGCCCCGTCAGATCACCATCTGCAACCAAAGCCCTGTGCGCCCACGTGCTGAACATCACGGACAAGGACCTGTC
TCCCCACACCTCCCCTTTCCAGGCCCAGCTCACAGATGACTCAGACATCTACTGGACGGCAGAGGTCAACGAGGAAG
GTGACACAGTGGTCTTGTCCCTGAAGAAGTTCCTGAAGCAGGATACATATGACGTGCACCTTTCTCTGTCTGACCAT
GGCAACAAAGAGCAGCTGACGGTGATCAGGGCCACTGTGTGCGACTGCCATGGCCATGTCGAAACCTGCCCTGGACC
CTGGAAAGGAGGTTTCATCCTCCCTGTGCTGGGGGCTGTCCTGGCTCTGCTGTTCCTCCTGCTGGTGCTGCTTTTGT
TGGTGAGAAAGAAGCGGAAGATCAAGGAGCCCCTCCTACTCCCAGAAGATGACACCCGTGACAACGTCTTCTACTAT
GGCGAAGAGGGGGGTGGCGAAGAGGACCAGGACTATGACATCACCCAGCTCCACCGAGGTCTGGAGGCCAGGCCGGA
GGTGGTTCTCCGCAATGACGTGGCACCAACCATCATCCCGACACCCATGTACCGTCCTAGGCCAGCCAACCCAGATG
AAATCGGCAACTTTATAATTGAGAACCTGAAGGCGGCTAACACAGACCCCACAGCCCCGCCCTACGACACCCTCTTG
GTGTTCGACTATGAGGGCAGCGGCTCCGACGCCGCGTCCCTGAGCTCCCTCACCTCCTCCGCCTCCGACCAAGACCA
AGATTACGATTATCTGAACGAGTGGGGCAGCCGCTTCAAGAAGCTGGCAGACATGTACGGTGGCGGGGAGGACGACT
AGGCGGCCTGCCTGCAGGGCTGGGGACCAAACGTCAGGCCACAGAGCATCTCCAAGGGGTCTCAGTTCCCCCTTCAG
CTGAGGACTTCGGAGCTTGTCAGGAAGTGGCCGTAGCAACTTGGCGGAGACAGGCTATGAGTCTGACGTTAGAGTGG
TTGCTTCCTTAGCCTTTCAGGATGGAGGAATGTGGGCAGTTTGACTTCAGCACTGAAAACCTCTCCACCTGGGCCAG
GGTTGCCTCAGAGGCCAAGTTTCCAGAAGCCTCTTACCTGCCGTAAAATGCTCAACCCTGTGTCCTGGGGCCTGGGCC
TGCTGTGACTGACCTACAGTGGACTTTCTCTCTGGAATGGAACCTTCTTAGGCCTCCTGGTGCAACTTAATTTTTTTT
TTTTAATGCTATCTTCAAAACGTTAGAGAAAGTTCTTCAAAAGTGCAGCCCAGAGCTGCTGGGCCCACTGGCCGTCC
TGCATTTCTGGTTTCCAGACCCCAATGCCTCCCATTCGGATGGATCTCTGCGTTTTTATACTGAGTGTGCCTAGGTT
GCCCCTTATTTTTTATTTTCCCTGTTGCGTTGCTATAGATGAAGGGTGAGGACAATCGTGTATATGTACTAGAACTT
TTTTATTAAAGAAA
```

FIGURE 68

```
GCGCCTTGCCCGCGGCCCCGCCGGGCCTCAGGCCCGGGGACCCTCAGCGCTACCTTTTGCCTCCGGCTCTGGGACTG
CTCGTGGCCGCGGGCGTGGCCGGCGCTGCGCTCTTGCTGGTCCACGTGCGCCGCCGTGGCCACTCCCAGGATGCTGG
GTCTCGCTTGCTGGCTGGGACCCCGGAGCCGTCAGTCCACGCACTCCCGGATGCACTCAACAACCTAAGGACGCAGG
AGGGTTCCGGGGATGGTCCGAGCTCGTCCGTAGATTGGAATCGCCCTGAAGATGTAGACCCTCAAGGGATTTATGTC
ATATCTGCTCCTTCCATCTACGCTCGGGAGGTAGCGACGCCCCTTTTCCCCCCGCTACACACTGGGCGCGCTGGGCA
GAGGCAGCACCTGCTTTTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXCGGGCGTGGCCGGCGCTGCGCTCTTGNTGGTCCANGTGCGCNGCCGTGGCCAATCNCAGGATGC
TGGGTCNCGCTTNCTGGTTNGGACCCCGGNGCCGTNAGTCCACNCACTCCNNGATGCACTCAACAACCTAAGGACGC
AGGAGGGTTCCCGGGGATGGTCCGAGCTCGTCCGTAGATTGGAATCGCCCTGAAGATGTAGACCCTCACAGGGAGTG
TATGTAATXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTCCGCACGTGGAGTCAGAG
CGTGGATTTTTGTANNTGCTNGGTGGTGCCCAGTCTCTGCCCCNGAGGCTTTGGANTTCAATCTTGAAGGGGTGTNC
TGGGGGAACTTTACTGTTGCAAGTTGTAAATAATGGTTNTTTATATCGTNTTTTTTCGTCACCCCATCTCTGATAGA
AACACCTNTAAAGGNTATTATTGTGAGTCAAAAAAA
```

FIGURE 69

GCGGCGACTGCGGCGACCGCGGGACGGCGAGAGGCACGCGGCGGGAGGGGACCGGAATCCGCAGCTCCGGCCGCGCC
ATGGACGGCAACGACAACGTGACCCTGCTCTTCGCCCCTCTGCTGCGGGACAACTACACCCTGGCGCCCAATGCCAG
CAGCCTGGGCCCCGGCACGGACCTCGCCCTCGCCCCTGCCTCCAGCGCCGGCCCCGGCCCTGGGCTCAGCCTCGGGC
CGGGTCCGAGCTTCGGCTTCAGCCCCGGCCCCACTCCGACCCCGGGAGCCCACGACCAGCGGCCTCGCGGGCGGCGCG
GCGAGCCACGGCCCCTTCCCCGTTCCCTCGGCCCTGGGCGCCCACGCGGCTCCCGTTCTGGGACACGCCGCTGAACCA
CGGGCTGAACGTGTTCGTGGGCGCCGCCCTGTGCATCACCATGCTGGGCCTGGGCTGCACGGTGGACGTGAACCACT
TCGGGGCGCACGTCCGTCGGCCCGTGGGCGCGCTGCTGGCAGCGCTCTGCCAGTTCGGCCTCCTGCCGCTGCTGGCC
TTCCTGCTGGCCCTCGCCTTCAAGCTGGACGAGGTGGCCGCCGTGGCCGGTGCTCCTGTGTGGCTGCTGTCCCGGCGG
CAATCTCTCCAATCTTATGTCCCTGCTGGTTGACGGCGACATGAACCTCAGCATCATCATGACCATCTCCTCCACGC
TTCTGGCCCTCGTCTTGATGCCCCTGTGCCTGTGGATCTACAGCTGGGCTTGGATCAACACCCCTATCGTGCAGTTA
CTACCCCTAGGGACCGTGACCCTGACTCTCTGCAGCACTCTCATACCTATCGGGTTGGGCGTCTTCATTCGCTACAA
ATACAGCCGGGTGGCTGACTACATTGTGAAGGTTTCCCTGTGGTCTCTGCTAGTGACTCTGGTGGTCCTTTTCATAA
TGACCGGCACTATGTTAGGACCTGAACTGCTGGCAAGTATCCCTGCAGCTGTTTATGTGATAGCAATTTTTATGCCT
TTGGCAGGCTACGCTTCAGGTTATGGTTTAGCTACTCTCTTCCATCTTCCACCCAACTGCAAGAGGACTGTATGTCT
GGAAACAGGTAGTCAGAATGTGCAGCTCTGTACAGCCATTCTAAAACTGGCCTTTCCACCGCAATTCATAGGAAGCA
TGTACATGTTTCCTTTGCTGTATGCACTTTTCCAGTCTGCAGAAGCGGGGATTTTTGTTTTAATCTATAAAATGTAT
GGAAGTGAAATGTTGCACAAGCGAGATCCTCTAGATGAAGATGAAGATACAGATATTTCTTATAAAAAACTAAAAGA
AGAGGAAATGGCAGACACTTCCTATGGCACAGTGAAAGCAGAAAATATAATAATGATGGAAACCGCTCAGACTTCTC
TCT**TAA**ATGTGGAGATACACAGGAGCTTCTATCTTGCTGAAATATTGCTTCATATTTATAGCCTGTGGTAGTGCACAT
GGTTAACATAAAAGATAACACTGGTTCACATCATACATGTAACAATTCTGATCTTTTTAAGGTTCACTGGTGTATTA
ACCAAACGTTGTCACAAATTACAAATCAATGCTGTAATATAATTTGCACCTGGAATGGCTAACGTGAAGCCTGAATT
AAATGTGGTTTTTAGTTTTTTACCATCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 70

CGGCCAGCACACCCCGGCACCTCCTCTGCGGCAGCTGCGCCTCGCAAGCGCAGTGCCGCAGCGCACGCCGGAGTGGC
TGTAGCTGCCTCGGCGCGGCTGCCGCCCTGCGCGGGCTGTGGGCTGCGGGCTGCGCCCCCGCTGCTGGCCAGCTCTG
CACGGCTGCGGGCTCTGCGGCGCCCGGTGCTCTGCAACGCTGCGGCGGGCGGCATGGGATAACGCGGCC**ATG**GTGCCG
CCGAGATCGCCTCCGCAGGATGAGGGAGTGGTGGGTCCAGGTGGGGCTGCTGGCCGTGCCCCTGCTTGCTGCGTACC
TGCACATCCCACCCCCTCAGCTCTCCCCTGCCCTTCACTCATGGAAGTCTTCAGGCAAGTTTTTCACTTACAAGGGA
CTGCGTATCTTCTACCAAGCACTCTGTGGGTGTGGTTGGAAGTCCAGAGATAGTTGTGCTTTTACACGGTTTTCCAAC
ATCCAGCTACGACTGGTACAAGATTTGGAAGGGTCTGACCTTGAGGTTTCATCGGGTGATTGCCCTTGATTTCTTAG
GCTTTGGCTTCAGTGACAAACCGAGACCACATCACTATTCCATATTTGAGCAGGCCAGCATCGTGGAAGCGCTTTTG
CGGCATCTGGGGCTCCAGAACCGCAGAATCAACCTTCTTTCTCATGACTATGGAGATATTGTTGCTCAGGAGCTTCT
CTACAGGTACAAGCAGAATCGATCTGGTCGGCATACCATAAAGAGTCTCTGTCTGTCAAATGGAGGTATCTTTCCTG
AGACTCACCGTCCACTCCTTCTCCAAAAGCTACTCAAAGATGGAGGTGTGCTGTCACCCATCCTCACACGACTGATG
AACTTCTTTGTATTCTCTCGAGGTCTCACCCCAGTCTTTGGGCCGTATACTCGGCCCTCTGAGAGTGAGCTGTGGGA
CATGTGGGCAGGGATCCGCAACAATGACGGGAACTTAGTCATTGACAGTCTCTTACAGTACATCAATCAGAGGAAGA
AGTTCAGAAGGCGCTGGGTGGGAGCTCTTGCCTCTGTAACTATCCCCATTCATTTTATCTATGGGCCATTGGATCCT
GTAAATCCCTATCCAGAGTTTTTGGAGCTGTACAGGAAAACGCTGCCGCGGTCCACAGTGTCGATTCTGGATGACCA
CATTAGCCACTATCCACAGCTAGAGGATCCCATGGGCTTCTTGAATGCATATATGGGCTTCATCAACTCCTTC**TGA**G
CTGGAAAGAGTAGCTTCCCTGTATTACCTCCCCTACTCCCTTATGTGTTGTGTATTCCACTTAGGAAGAAATGCCCA
AAAGAGGTCCTGGCCATCAAACATAATTCTCTCACAAAGTCCACTTTACTCAAATTGGTGAACAGTGTATAGGAAGA
AGCCAGCAGGAGCTCTGACTAAGGTTGACATAATAGTCCACCTCCCATTACTTTGATATCTGATCAAATGTATAGAC
TTGGCTTTGTTTTTTGTGCTATTAGGAAATTCTGATGAGCATTACTATTCACTGATGCAGAAAGACGTTCTTTTGCA
TAAAAGACTTTTTTTTAACACTTTGGACTTCTCTGAAATATTTAGAAGTGCTAATTTCTGGCCCCACCCCCCAACAGGA
ATTCTATAGTAAGGAGGAGGAGAAGGGGGGCTCCTTCCCTCTCCTCGAATGACGTTATGGGCACATGCCTTTTAAAA
GTTCTTTAAGCAACACAGAGCTGAGTCCTCTTTGTCATACCTTTGGATTTAGTGTTTCATCAGCTGTTTTTAGTTAT
AAACATTTTGTTAAAATAGATATTGGTTTAAATGATACAGTATTTTAGGTATGATTTAAGACTATGATTTACCTATA
CATTATATATTTTATAAAGATACTAAACCAGCATACCCTTACTCTGCCAGAGTAGTGAAGCTAATTAAACACGTT
TGGTTTCTGAATAAATTGAACTAAATCCAAACTATTTCCTAAAATCACAGGACATTAAGGACCAATAGCATCTGTGC
CAGAGATGTACTGTTATTAGCTGGGAAGACCAATTCTAACAGCAAATAACAGTCTGAGACTCCTCATACCTCAGTGG
TTAGAAGCATGTCTCTCTTGAGCTACAGTAGAGGGGAAGGGATTGTTGTGTAGTCAAGTCACCATGCTGAATGTACA
CTGATTCCTTTATGATGACTGCTTAACTCCCCACTGCCTGTCCCAGAGAGGCTTTCCAATGTAGCTCAGTAATTCCT
GTTACTTTACAGACAGGAAAGTTCCAGAAACTTTAAGAACAAACTCTGAAAGACCTATGAGCAAATGGTGCTGAATA
CTTTTTTTTTAAAGCCACATTTCATTGTCTTAGTCAAAGCAGGATTATTAAGTGATTATTTAAAATTCGTTTTTTTA
AATTAGCAACTTCAAGTATAACAACTTTGAAACTGGAATAAGTGTTTATTTTCTATTAATAAAAATGAATTGTGACA
AAAAAAAAACCG

FIGURE 71A

```
CGGGTGGACAAGTCACCTGTCTAACTGTGGTGAGAGCAACAGGCCTCCAAAGGAAAGAAGTTGTTTCCGAGTTTGTG
ACTGGCACAGTGACCTCTTTCAGTGGGAGGTTTCTGACTGGCACCACTGTGTGCTTGTTCCTTACGCTCGCGGTGAA
GTCAAGCCTCGGACTGCAGAGTGTGTGACGGCTCAGCATGGACTGCAGCACCGGATGGTGCGCTGCATTCAGAAGCT
GAACCGAACTGTGGTTGCAAATGAAATATGCGAACACTTTGCCCTTCAGCCTCCTACAGAACAGGCTTGCCTCATTC
CTTGTCCCCGGGATTGTGTAGTATCTGAGTTCTTACCATGGTCCAACTGTAGCAAGGGATGTGGGAAGAAATTGCAG
CATAGAACTCGCGCGGTCATAGCTCCCCCTCTCTTTGGTGGTTTGCAATGTCCAAATCTGACTGAGTCAAGAGCCTG
TGATGCTCCCATTTCCTGTCCTCTGGGGAGAGGAATATACATTTAGCCTTAAGGTTGGACCATGGAGTAAATGCA
GACTGCCTCATCTTAAAGAAATTAATCCAAGCGGAAGAACTGTTCTGGATTTTAACTCTGATTCAAATGAGCGAGTC
ACCTTTAAACATCAAAGTTACAAAGCACATCATCATTCGAAGTCTTGGGCAATAGAGATAGGTTATCAAACCCGGCA
GGTTTCGTGTACAAGAAGTGATGGACAAAATGCTATGTTAAGCCTTTGCCTTCAAGATTCCTTCCCATTGACTGTTC
AGTCCTGCATCATGCCCAAAGACTGTGAAACCTCCCAGTGGTCCTCCTGGAGCCCCTGCTCCAAGACATGCCGTTCA
GGGAGTCTCTTGCCAGGATTTAGGAGCAGGAGCCGGAACGTGAAGCACATGGCTATTGGAGGTGGAAAGGAGTGTCC
TGAACTTCTTGAGAAAGAGGCCTGCATTGTTGAAGGAGAACTTCTGCAGCAATGTCCCAGGTATTCCTGGAGAACTT
CTGAATGGAAAGAATGCCAAGTCTCTCTCCTCCTCGAGCAGCAGGATCCCCACTGGCATGTGACGGGACCCGTGTGT
GGCGGTGGGATCCAGACCCGGGAGGTGTACTGTGCCCAGAGCGTACCAGAGCTGCCGCACTGCAGGGCCAAGGAAGT
CTCTAGACCTGTGGAAAAGGCATTATGTGTGGGACCCGCCCCGTTGCCCTCTCAGCTCTGCAATATCCCTTGCTCTA
CGGACTGCATAGTATCTTCCTGGTCAGCCTGGGGCCTGTGCATCCATGAAAACTGTCATGAACCTCAGGGGAAAAAA
GGATTTAGAACGAGGCAGCGCCATGTCCTCATGGAATCTACAGGGCCTGCAGGGCATTGCCCTCATTTGGTGGAGTC
TGTTCCTTGTGAGGATCCAATGTGCTACCGATGGCTGGCATCAGAAGGGATCTGTTTCCCTGATCATGGAAAATGTG
GCCTGGGACATCGTATTCTGAAGGCCGTCTGCCAGAATGACCGCGGAGAAGATGTATCAGGGAGTCTTTGCCCAGTT
CCCCCTCCTCCTGAGAGGAAGTCTTGTGAAATTCCCTGCCGAATGGACTGTGTGCTGAGCGAGTGGACGGAGTGGTC
ATCCTGTTCCCAGTCCTGTTCAAATAAAAACTCAGATGGGAAACAGACCAGGTCAAGAACTATCCTGGCACTGGCTG
GGGAAGGTGGAAAGCCATGTCCCCCTAGTCAGGCTCTCCAAGAGCATCGTTTGTGTAATGACCATTCCTGTATGCAG
CTTCACTGGGAGACATCGCCTTGGGGCCCTTGTTCTGAGGACACATTGGTAACTGCCCTTAATGCAACCATTGGCTG
GAATGGAGAAGCCACGTGTGGTGTAGGCATTCAGACTCGGAGAGTCTTCTGTGTCAAGAGTCACGTGGGACAAGTAA
TGACCAAAAGATGTCCAGATTCTACTCGACCTGAAACTGTGCGCCCCTGTTTTCTCCCATGCAAAAAAGACTGTATT
GTGACTGCTTTCAGTGAGTGGACACCCTGCCCAAGGATGTGCCAAGCAGGAAATGCCACAGTAAAACAGTCTCGATA
CAGAATCATCATCCAAGAAGCAGCCAATGGAGGCCAGGAATGCCCAGATACCTTATATGAGGAGAGAGAGAGTGTGAAG
ATGTTTCCTTGTGTCCTGTATATCGGTGGAAGCCACAGAAATGGAGCCCTTGCATCTTAGTGCCAGAGTCTGTCTGG
CAGGGAATAACGGGCAGCAGTGAAGCCTGTGGAAAGGGGTTACAAACAAGAGCTGTCTCATGCATCTCTGATGACAA
CCGGTCAGCAGAAATGATGGAATGCCTCAAGCAGACAAACGGCATGCCTCTCCTTGTGCAAGAATGCACAGTCCCAT
GTCGAGAAGACTGCACCTTCACTGCTTGGTCCAAGTTTACGCCCTGCTCCACGAACTGTGAAGCCACAAAAAGTAGG
CGGCGACAGCTCACAGGGAAAAGCAGAAAGAAGGAGAAATGCCAGGATTCTGACCTTTACCCTCTAGTGGAGACTGA
ACTATGTCCTTGTGATGAATTTATATCCCAACCTTATGGAAACTGGTCAGATTGCATTCTTCCAGAAGGCAGAAGGG
AGCCTCACCGAGGACTGCGGGTACAAGCAGACAGCAAGAAATGTGGAGAAGGCCTGCGCTTTCGAGCAGTAGCCTGT
TCTGATAAAAATGGAAGACCTGTTGACCCCTCCTTCTGCAGCAGCTCTGGTTACATTCAAGAAAAATGTGTCATTCC
CTGCCCATTTGATTGCAAGTTAAGCGATTGGTCTAGTTGGGGGTCTTGCAGTTCATCTTGTGGAATTGGAGTGAGAA
TTCGATCCAAATGGCTAAAAGAAAAACCTTACAATGGAGGACGACCATGTCCCAAACTGGATCTCAAGAATCAGGCT
CAGGTACATGAGGCAGTCCCATGTTACAGTGAGTGCAATCAGTATTCCTGGGTTGTAGAACACTGGTCTTCATGCAA
AATCAACAATGAGCTGAGGTCCCTGCGCTGTGGAGGAGGAACACAATCTAGGAAAATCAGATGTGTGAATACTGCGG
ATGGTGAAGGTGGAGCAGTGGATAGCAACCTGTGCAACCAGGATGAAATTCCCCCAGAAACCCAGTCCTGTTCTCTT
ATGTGTCCCAATGAGTGTGTCATGTCTGAGTGGGGACTTTGGAGCAAATGCCCACAGTCATGCGATCCCCACACAAT
GCAGAGAAGAACTCGCCACCTGCTAAGACCATCACTGAACTCAAGGACTTGTGCTGAAGACTCACAGGTGCAGCCTT
GCCTCCTGAATGAAAATTGCTTCCAGTTCCAGTACAATCTAACAGAGTGGAGCACATGCCAGCTGAGTGAAAACGCA
CCCTGTGGTCAAGGCGTCAGGACCCGCCTGCTAAGCTGTGTGTGCAGTGATGGCAAGCCAGTCAGCATGGACCAATG
TGAGCAGCATAATTTGGAGAAGCCCCAGAGAATGAGCATTCCCTGCTTGGTGGAATGCGTGTGTCAACTGTCAGCTCT
CAGGGTGGACGGCTTGGACAGAGTGTTCACAGACGTGTGGCCATGGAGGTCGAATGAGCCGGACTCGATTTATCATT
ATGCCAACCCAAGGAGAAGGACGGCCATGCCCCACAGAGCTTACCCAGGAGAAAACCTGCCCAGTGACCCCCTGCTA
CAGCTGGGTCCTTGGCAACTGGTCTGCATGTAAATTGGAGGGTGGAGACTGTGGGGAAGGAGTTCAGATCCGGCAGCT
TTTCCTGCATGGTCCACAGTGGTTCAATATCTCATGCAGCTGGACGTGTCGAGGATGCACTGTGTGGAGAAATGCCC
TTTCAGGACAGCATCCTGAAGCAGCTGTGTTCTGTGCCTTGCCCAGGAGACTGCCATTTAACAGAATGGTCAGAGTG
GAGCACATGTGAATTAACCTGCATTGATGGGAAGAAGCTTTGAGACTGTGGGCCGCCAGTCTAGATCAAG
```

FIGURE 71B

GACTTTTATAATTCAGTCTTTTGAGAACCAAGACAGCTGCCCCCAACAGGTTCTAGAAACACGCCCTTGTACAGGAG
GCAAATGTTATCACTACACATGGAAAGCAAGTCTTTGGAACAATAACGAACGAACTGTATGGTGCCAGCGTTCAGAT
GGCGTTAATGTCACAGGAGGCTGCTCCCCTCAGGCCCGTCCTGCTGCCATTCGGCAGTGCATTCCAGCCTGCAGAAA
ACCTTTCTCCTACTGTACACAGGGTGGAGTCTGTGGTTGTGAGAAGGGCTATACAGAGATAATGAGATCAAATGGTT
TCCTGGATTACTGCATGAAAGTACCAGGCTCAGAGGATAAAAAAGCTGATGTGAAAAACCTTTCTGGGAAAAACAGA
CCTGTGAATTCAAAAATACATGATATTTTTAAAGGATGGTCTCTTCAACCACTTGATCCAGATGGCCGAGTAAAAAT
TTGGGTTTATGGCGTTTCAGGTGGCGCTTTTCTCATCATGATTTTCCTAATATTTACTTCCTACCTTGTTGCAAGA
AGCCAAAACCACATCAAAGCACACCTCCCCAACAGAAGCCTCTGACCTTAGCCTACGATGGAGACTTAGACATG**TAA**
TCTGAAAAAGAAATCCAAATGTAGACATCAACTGCCTTAACCGCTTTCTCTTTTGTAGCTCTCAGACTTCTCAGTTT
TTTGAGGAATCTCAAGATGTGATATATTGGGCAGAATACAAATATTGCAAAAGTAATATTGCCTCAACTTCATTTGG
ACATGGAGTCAAGGATTATTAGGTCTGCCATTTTGTTTTCAAGTTGTTTGTGGGTGTGTTTTATTTTTTTGGTTTTC
CCAAGGGACCTGAAAACCCTTCTCTTCCTGTTTGGAACTGGGAGGAAGAAAACATGATGGAATTCCCACAGACTTGA
GTAAACTTGATCTTCAGCAGCATAATGACAATCCAGAGGAAAATACAGTCAAGGCATTTACTCTAAATGACGAGTCT
GACACTGCATTGTTACGCACTATATTAGTGCAAAGTCTTTATTCTTCCCATACTTCAACACTGAGTTTTCTAGAGTT
TACTTTGGTTTAAAGACTTTCAAATTGGATTGCCTATTTTTCATGAACACAGAGAGAATGGATTACCATTTCAGAAA
TTCTCTGAGTTTTTAACCTTTAAATATTGTATTTTGTTTTGTTTGTAGCCAGGGGATGATGGCGCTTCATGGGTTGCAGCT
ACTGAAAATAGCAGCGTGTGTGTAATTGCTGGACTAGATGAAAGCTAGGTCATTTCTGAAGGGAATGTGTACTGAAT
GTTAGAGTGTACAAATGAAATATGTGGTTAAATTGGAGAATGAGGTAGATTATTTGATTACTAAAACTGTATTTTAA
CAAAAACTTATCCATGTAGATATAGCATTAACCACACACAGTTGTAATTCAGTTTAATGATGACAAACTCTGCTTTT
GTAATTTCAATTTTCTTATCTGAATATTTATAAATTCTTCTTTCAAATTTAATTATCTGACCTCATTTAATATACAT
CAAACACCGATCCTGTTTGTACAAAGTCTTGCTTTTATAAGGTTTCAATAATATCTAAAACAACACATTAAAAAGCT
GAGACCATTTTATGAAGATAATTGTTTGTAATCATAGGTGTTGAAAGTAAAAAGGTGCCATCTTGTGGTATTGACTT
GTATTTATAACAAATAAACTGCTCAAGAGACTGC

FIGURE 72A

CCCACCTCGGCCACACTGCTTCGCTCCCCCCTCCCCACCTTCAGCTGGCACCGCAAACAAGCCTTACCTCGTTGCAT
CTGCGAGGAGAGGTAGCAACAGCGAGCCTAGCCAGCCAGAGGCGGTGGAGAGGAGGAAGGGGCGGGGTGGGGGGGCA
GAGAGCGGGCCGAGGCCGCCCTTGGTGGGGGTAGCGGGGGCAGAGCTGCCGAGCAGACCCGGCAGCCGCCCTCCTCC
GCCCCCACCCTCCAGAAGCACTCTTGCCTGAGAACCGAATTATTCCACTAGTATTATTTCATTTTTTATTACTCCCC
CTCCTCACTCCCCAGCCGCCCCCCACCCCCACGCCGGCCTCGCCTCTGGTTGCATGGCAGCGCTGCCCGGGCGCGGGG
GCTCAGGGCTGGCCCCCAGGGAAGGGGGAGGAGGAGGAGGATCATGAAGGCCGGAGTCGCGACCGCGCCGGACGGCG
GGCAGCAGCCAGAGGACGAGCCGGAGCAGCCCCCGCCCCGGAGACACCCGGACGCCGAGCAGCAGTCGCCTCCGCCG
CCGCAGCAGCAGCTGCGGGCGCGGGCCGACCCGGAGCCGGAGGAGGAGGGGCGACGCGGACCCAGAGGAGGAAGAGGA
GGAGAAGGCGAGGCCACTCGCCCGCCCGGACCTGCCCTTCTCTCGGCTCTTGCCCCTCGACCGAAGGGACCTTTGAT
GGAACCGAGGGGAGGGCGCCACGGATTTGCCGACTGCAGCAGGGGTGGGCTGGGGGCTGAGATAATGTAACCACTCC
TTTCTCCTGTTCTCTCCCACACGCCCCTCTCCTCTACCCCTATTCTCTGCTCCACTGCCCTCTCACCCCGGTACACA
CACCCTTCCTCTAGCCAGGATCTTCATGCTCAGGAAGGAGGCGCCTCTGCAAGGGTTAAACGATCTTTTCTTTTTCT
CCCATCCTTTTTCCTTCCCAAACCTCTATTTTACCTCCCTTCTCCTAATTGGCTTTCCCCTCTTCGGTGTAACCCTT
TGGCTGCGGAGGCAAAGCACAAGCCCCTAGCCCAGTTTCACCTGCAACACCCCTCCCCCACCCAACTGCTCTCTTAA
AAGCAACTCTGGTGCTTCTGGGGGTTAATTGCCCCAGTTTTCTGCCCAGGAGAATTAAAACTTCTCCCAATCTTCTC
TCCTCCCCTACCTTGACTCCCCCAACCCCTACCACCTGAGAAAAACGATCTTTTCTCTCTCACACATGCAGTCCTCA
ATTCTTCATTGAGCTAGTTTTCTCTAAGCCCAGCTCAATCCACTCCAGATTTGATTTACAATTGTCCCCACCCTTTT
ATATAAAAGAAAGATTTCTCACTGCGTAGGAATTTGAGAAGAACCCAATAATCCTTTCCTGGGGAACTTTTAAACAA
TTCGACATTGATTTAAACAATTCGACAGAGGCTCTAGTGGCCCCTCCACCACGCTTCTTCAATCCCTCTGCTCCTAC
CAGTGTCCTCAGGTCAAAAGCAGAAAGGAGACACACTGAACAAAGTTGGAGGTTGGGGTGGGTGTGTGAGGGCAAGA
AAAACTTTTTTGTTATTGGGCTTTCCAGGTGGAGTTCAGAACCAGTGACTCACACTTCTCAGTCCTGGGAGCAATTT
ATTTGCTACTTGGAGGGGTTGTAAGAAAAGCCAGTGAGAAAGCAGACTCCCCCCACAACACAGATCCACTGTGGACC
CCCAAAACCTGTCCTGTCCCCCTCTTTTAAGACTCCAGCCACCCCTCTTGGGCTCTCTACTTCCACGGGGCACATGC
TGATGCCCCTGTGTGGGCTGCTCTGGTGGTGGTGGTGGTGCTGCTGCTCCGGCTGGTACTGCTATGGATTGTGTGCCCCA
GCCCCCCAGATGTTGCGGCCACCAGGGTCTCCTCAAGTGCCGCTGCCGCATGCTCTTCAATGACCTGAAGGTTTTCTT
ACTGCGGCGCCCTCCTCAAGCGCCCCTGCCCATGCACGGCGACCCCCAGCCCCCCGGTTTGGCGGCCAACAACACCC
TTCCGGCTCTGGGCGCCGGGGGGTGGGCAGGCTGGAGGGGCCCCCGAGAAGTGGTGGGCAGGGAGCCCCCTCCTGTG
CCACCTCCACCCCCCTTGCCACCTTCTTCTGTGGAAGATGACTGGGGTGGCCCAGCCACAGAGCCACCTGCCTCGCT
GCTCAGCAGTGCCTCCTCAGATGACTTCTGTAAGGAGAAGACCGAGGATCGCTACTCACTGGGCAGCAGCTTGGACA
GTGGTATGAGGACCCCACTCTGCCGCATCTGCTTCCAGGGGCCAGAACAGGGGGAGCTGCTGAGCCCATGCCGCTGT
GATGGCTCGGTCAAGTGCACACCACCAGCCTTGCCTCATCAAGTGGATCAGCGACGGGGGCTGCTGGAGCTGCGAGCT
GTGCTACTACAAGTACCACGTCATCGCCATAAGCACCAAAAAATCCTCTGCAGTGGCAGGCCATCTCTCTGACGGTCA
TTGAGAAGGTTCAGGTTGCAGCCGCCATCCTGGGCTCCCTCTTCCTCATCGCCAGTATTTCTTGGCTCATCTGGTCA
ACTTTCAGCCCCTCGGCAAGATGGCAGCGCCAAGACCTTCTCTTCCAGATCTGCTACGGGATGTATGGCTTCATGGA
CGTGGTGTGCATAGGTCTCATCATCCATGAAGGACCCTCGGTGTACCGCATCTTTAAACGGTGGCAGGGAGCCCCCTGT
AGCAGTGGAAAGTGCTGAACTATGACAAGACAAAAGACCTGGAGGATCAAAAGGCAGGAGGCAGGACCAACCCCCGG
ACCTCCTCATCCACCCAGGCCAATATCCCCTCCTCGGAAGAGGAGACCGCAGGCACCCCTGCCCCTGAGCAGGGCCC
TGCCCAGGCTGCCGGCCACCCCTCAGGCCCTCTGTCCCATCACCACTGTGCTTATACCATCCTGCACATCCTGAGTC
ACTTGAGACCTCATGAACAGCGAAGTCCCCCAGGCAGCAGCCGAGAGCTGGTCATGAGAGTCACGACAGT**GTGA**GAG
CAGAGGCCCGAAGGAAGGCCATGACCACCACTGAGGGCCCGGAGCAGGGTGGGGAGGTGCAGTGGCACCCCCGGAG
CCAACAGAGGGAGCAGGCAGAGGGTGGGGACCTGGCGGGAGCCCTGGGGTAGTGTCAGAGCGGGAGTGAGGCTGGT
GCAGGAGCAGTTCTGCTATTTCCAATCAGTCAATGCCACTCTTCCACAACAACAATGAAAACCAACACCAACTCAACA
ACAAAGTGCAATACAGGCTGAACCTGGCCCAACAGAAAAACCCTGCCCCAATGCACCTGCAGGCAAGGTACCCGAAG
AAGCAGAGGCTGAGGGCAGGCAAAGCCTGTGTGACTGTGGCAGTGCCCGGAGGCCAAGGGGGGCCAAGAGGAAAAGCAT
CTGTGGTCTGCCCTGCTCTCACCCTGTTTGGTTTTGTTTCTCCTGGGGCTGTGTTCTGCAGGCAGCCAGAAAAGGAG
GAGGCACGGGTGAGCTGGCAGGGACACACTGCCTTTGGGGGCTCCTGGGCTCATTTGGATGAGCAAGATTCGCTGACA
AATGGCTGTGGGGATGGTGGGGTGGATGGTCAGGGAGGGATCCTCAGGGAGGGATATGCTGGTGTGAGCAGCCAGAG
GGAGAGTGTGTCTCCTTCCTGAAGGAACTTCCAAATGGAACTCCCGATTTCAGGTGGGCTAAAAGAGGGCTTAGGTT
TGGAAAAGGGTGTCCTTCTGTGCCCTTGTTAATTTATTTTATAGTGATTTGGTTCAAAGATGTTTACAGGACACACA
CACACACACACACACACACACACACACACACACCCTAGAGAAAAGTACAGATTTCCAGTGGATATTTCAAGCAC
AGTTCTGCTGCTGTGTGGCTTCAGCTTTGGAAGCTGTCAATCCCGGAGCAACTTTCCCAACTACCCAACCCCACCATGG
CCAGGACATGTGCAATGCCAGCCCCTTCTTGTCTTGGCACATGCACAGACCCAGTCCCCTCACGGTAGG

FIGURE 72B

GCACCCCTGACCTACGGGCTTCCAAGAGAGCAGCTGCAGTGGTTGGGAGGAGCTTGACCAGTGTGCCCCAAGGAGTG
GAGTAGAGCCCAATCTAAGTATTCCTTGCTGCTTGGAACCCTCCCTGTTTGGAACCCTCCCCAAAGAGGCAGTCAGG
CTGATGCTCAGTGCTTTGTGCTCCCTGCTCCTTCCCGCGTAGCCAGGTGGGCCCAAGGGTGCCTGGCAGGGAGCACT
ACCCCTGGACCCCTCCTGCTCGCTCTGGGGACCCTGCCAGGGAAGGCCACTGGGTGTTCACCTGCAAAGTTTCTGGT
TGTCACTGCACAGTGGTCGCGTCATCCATGGGTATTAAAAGGACACTGTCAAGTACTTTTTTAAACTAGTTTTTAGG
GTTTTTTAAAACTCTCTGTTGTTTGTAATATTCTCTTAAAAGCTTGAAAATAAAACTTCTTTCCCTACC

FIGURE 73

```
CAGGAGCAAGGGGACAAGATGATGGAGGAATACAGCCTGGAGAAAAATGAGAGGGCCTGCATTGATTTCGCCATCAG
CGCCAAGCCGCTGACCCGACACATGCCGCAGAACAAGCAGAGCTTCCAGTACCGCATGTGGCAGTTCGTGGTGTCTC
CGCCTTTCGAGTACACGATCATGGCCATGATCGCCCTCAACACCATCGTGCTTATGATGAAGTTCTATGGGGCTTCT
GTTGCTTATGAAAATGCCCTGCGGGTGTTCAACATCGTCTTCACCTCCCTCTTCTCTCTGGAATGTGTGCTGAAAGT
CATGGCTTTTGGGATTCTGAATTATTTCCGCGATGCCTGGAACATCTTCGACTTTGTGACTGTTCTGGGCAGCATCA
CCGATATCCTCGTGACTGAGTTTGGGAATAACTTCATCAACCTGAGCTTTCTCCGCCTCTTCCGAGCTGCCCGGCTC
ATCAAACTTCTCCGTCAGGGTTACACCATCCGCATTCTTCTCTGGACCTTTGTGCAGTCCTTCAAGGCCCTGCCTTA
TGTCTGTCTGCTGATCGCCATGCTCTTCTTCATCTATGCCATCATTGGGATGCAGGTGTTTGGTAACATTGGCATCG
ACGTGGAGGACGAGGACAGTGATGAAGATGAGTTCCAAATCACTGAGCACAATAACTTCCGGACCTTCTTCCAGGCC
CTCATGCTTTCTCTTCCGGAGTGCCACCGGGGAAGCTTGGCACAACATCATGCTTTCCTGCCTCAGCGGGAAACCGTG
TGATAAGAACTCTGGCATCCTGACTCGAGAGTGTGGCAATGAATTTGCTTATTTTTACTTTGTTTCCTTCATCTTCC
TCTGCTCGTTTCTGATGCTGAATCTCTTTGTCGCCGTCATCATGGACAACTTTGAGTACCTCACCCGAGACTCCTCC
ATCCTGGGCCCCCACCACCTGGATGAGTACGTGCGTGTCTGGGCCGAGTATGACCCCGCAGCTTGCGGTCGGATTCA
TTATAAGGATATGTACAGTTTATTACGAGTAATATCTCCCCCTCTCGGCTTAGGCAAGAAATGTCCTCATAGGGTTG
CTTGCAAGCGGCTTCTGCGGATGGACCTGCCCGTCGCAGATGACAACACCGTCCACTTCAATTCCACCCTCATGGCT
CTGATCCGCACAGCCCTGGACATCAAGATTGCCAAGGGAGGAGCCGACAAACAGCAGATGGACGCTGAGCTGCGGAA
GGAGATGATGGCGATTTGGCCCAATCTGTCCCAGAAGACGCTAGACCTGCTGGTCACACCTCACAAGTCCACGGACC
TCACCGTGGGGAAGATCTACGCAGCCATGATGATCATGGAGTACTACCGGCAGAGCAAGGCCAAGAAGCTGCAGGCC
ATGCGCGAGGAGCAGGACCGGACACCCCTCATGTTCCAGCGCATGGAGCCCCGTCCCCAACGCAGGAAGGGGGACC
TGGCCAGAACGCCCTCCCCTCCACCCAGCTGGACCCAGGAGGAGCCCTGATGGCTCACGAAAGCGGCCTCAAGGAGA
GCCCGTCCTGGGTGACCCAGCGTGCCCAGGAGATGTTCCAGAAGACGGGCACATGGAGTCCGGAACAAGGCCCCCCT
ACCGACATGCCCAACAGCCAACGTAACTCTCAGTCCGTGGAGATGCGAGAGATGGGCAGAGATGGCTACTCCGACAG
CGAGCACTACCTCCCCATGGGAAGGCCAGGGCCGGGCTGCCTCCATGCCCCGCCTCCCTGCAGAGAACCAGACCATCT
CAGACACCAGCCCCATGAAGCGTTCAGCCTCCGTGCTGGGCCCCAAGGCCCGACGCCTGGACGATTACTCGCTGGAG
CGGGTCCCGCCCGAGGAGAACCAGCGGCCACCACCAGCGGCGCCGCGACCGCAGCCACCGCGCCTCTGAGCGCTCCCT
GGGCCGCTACACCGATGTGGACACAGGCTTGGGGACAGACCTGAGCATGACCACCCAATCCGGGGACCTGCCGTCGA
AGGAGCGGGACCAGGAGCGGGGGCCGGCCCAAGGATCGGAAGCATCGACAGCACCACCACCACCCACCACCACCAC
CACCCCCCGCCCCCCGACAAGGACCGCTATGCCCAGGAACGGCCGGACCACGGCCGGGCACGGGCTCGGGACCAGCG
CTGGTCCCGCTCGCCCAGCGGCCCGAGGAGCCACTGGCCGCACCGGCAGGGCAGTAGTTCCGTAAGTGGAAGCCCAG
CCCCCTCAACATCTGGTACCAGCACTCCGCGGCGGGGCCGCGCCAGCTCCCCCCAGACCCCCTCCACCCCCCGGCCA
CACGTGTCCTATTCCCCTGTGATCCGTAAGGCCGGCGGCTCGGACGCAGCAGCAGCAGCAGCAGCAGCACA
GCAGCAGCAGGCGGTGGCCAGGCCGGGCCGGGCGGCACCAGCGGCCCTCGGAGGTACCCAGGCCCCACGGCCGAGC
CTCTGGCCGGAGATCGGCCGCCCACGGGGGGCCACAGGCAGCGGCCGCTGCCCCAGGATGGAGAGGCGGGTCCCAGGC
CCGGCCCCGGAGTCCCCCCAGGGCCTGTCGACACGGCGGGGCCCGGTGGCCGGCATCTGGCCCGCACGTGTCCGA
GGGGCCCCCGGGTCCCCGGCACCATGGCTACTACCGGGGCTCCGACTACGACGAGGTCCGCTCTGGGGCGCCACCGGGCGC
GCGGCGAGGAGGCCATGGCCGGGGCCTACGACGCGCCACCCCCCGTACGACACGCGTGCCTCGGGCGCCCACCGGGCGC
TCGCCCAGGACTCCCCGGGCCTCGGGGCCCGGCCTGCGCCTCGCCTTCTCGGCCACGGCCGGCGACTCCCCAACGGCTA
CTACCGGCGCACGGACTGGCCAGGCCCCGCGGGGCCGGGCTCCAGGAAGGGCCTGCACGAACCCTACAGCGAGAGTG
ACGATGATTGGTGC**TAA**GCCCGGGCGAGGTGGCGCCCGCCGGCCCCCACGCACCCCAGCGACCCCAGTGCCCTACCCTTACCCGA
GGAGCCGCGCAGAGGCCGCGGGGGGCCCAGCACGAGGGCCCGGGAGAGGGCCAGCCGGGAGACCCCAGACTCTGGAG
AGGCCAGGGCTGGGCCACAAGGGTGTCCCGCAGAGACCCTCGGCCCAAAAGAGACCCTCCTGGGCAGCCACGGCGCCT
CCCAACCAGCCCCGATCCCCCCACCCACGACAGGGGCTCTCGGGTGGGAGGCAGGGAGGCAGACAAAACCACACAGCCA
AGGGATTTGAATTAACTCAGCCATTTTTGGAGAACTTTGGGGAACATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACATT
TTTAAAAGAAAAAACGGGGAGAAAAAAAATAGCTTCTATTGATGAGTTTTATCATCTCAATTGAATCTTTCCTTTCCC
TGATGAAGACAGCTGGTGGCCGAGTGCGGCAAAGAAGCCAGAAGGAACCAGAATCCCAGTGCCCTACACCCACCACC
AGACACACTCACACCCACACACGTTCTCAGACACACACAAGAGTGCTTGCCGGTTATACCAAACCCTACTATTACTG
CCTGCAGAAATCAATTTAAAAAAATAATAATAACAATAAACAATTTTAAAAAGGACAAAAAAATTAATGATTGAGAA
AAGAGGCATTTTTTTCTGACATTTGGTCCTGCTTGAAACAACAAAAGAAGAAGAAAAAACCCACCATCACCACCGATT
CCTTTGCTTCTTTTTTCCTTTTTTCCTACCTTGTTTGAAAACCGTGGGCTTGGGACTGTGAA
```

FIGURE 74A

```
CTGCCCCTGGGTCTCTGCGCCTTTGCATGAGACTTTACGGTAAGCCGCTCCTCCCGCGCCCCCGCCCCCAGCCCCGC
TCGGCGATCCCCGGCGCCGTCGCCAGGCGCTGGCCGTGGTGCTGATTCTGTCAGGCGCTGGCGGCGGCAGCGGCGGT
GACGGCTGCGGCCCCGCTCCCTCTACCCGGCCGGACCCGGCTCTGCCCCCGCGCCCAAGCCCCACCAAGCCCCCCGC
CCTCCCGCCGCGGTCCCAGCCCAGGCGGCGGCCGCAACCAGCACCATGCGCCCGGTAGCCCTGCTGCTCCTGCCCTC
GCTGCTGGCGCTCCTGGCTCACGGACTCTCTTTAGAGGCCCCAACCGTGGGGAAAGGACAAGCCCCAGGCATCGAGG
AGACAGATGGCGAGCTGACAGCAGCCCCCACACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTCACAACAGCCCCC
ACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATTCCTACACGAGGGGCTGGAAAAGGGAGATGAGGAGCT
GAGGCCAGCACTGTCCTTTCAGCCTGACCCACCTGCACCCTTCACCCCAAGTCCCCTTCCCCGCCTGGCCAACCAGG
ACAGCCGCCCTGTCTTTACCAGCCCCACTCCAGCCATGGCTGCGGTACCCACTCAGCCCCAGTCCAAGGAGGGACCC
TGGAGTCCGGATCCGGAGTCAGAGTCCCCTATGCTTCGAATCACAGCTCCCCTACCTCCAGGGCCCAGCATGGCAGT
GCCCACCCTAGGCCCAGGGGAGATAGCCAGCACTACACCCCCCAGCAGAGCCTGGACACCAACCCAAGAGGGTCCTG
GAGACATGGGAAGGCCGTGGGTTGCAGAGGTTGTGTCCCAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCTCC
TCCACAGCTTCAGGAGATGATGAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCAGTCCAGAC
ACCAGGCCCTTGTAGCTGGAATTTCTCAGGCCCACAGAGGGCTCTCTGGACTCCCCTACAGACCTCAGCTCCCCCACTG
ATGTTGGCCTGGACTGCTTCTTCTACATCTCTGTCTACCCTGGCTATGGCGTGGAAATCAAGGTCAAGAATATCAGC
CTCCGGGAAGGGGAGACAGTGACTGTGGAAGGCCTGGGGGGGCCTGACCCACTGCCCCTGGCCAACCAGTCTTTCCT
GCTGCGGGGCCAAGTCATCCGCAGCCCCACCCACCAAGCGGCCCTGAGGTTCCAGAGCCTCCCGCCACCGGCTGGCC
CTGGCACCTTCCATTTCCATTACCAAGCCTATCTCCTGAGCTGCCACTTTCCCCGTCGTCCAGCTTATGGAGATGTG
ACTGTCACCAGCCTCCACCCAGGGGGTAGTGCCCGCTTCCATTGTGCCACTGGCTACCAGCTGAAGGGCGCCAGGCA
TCTCACCTGTCTCAATGCCACCCAGCCCTTCTGGGATTCAAAGGAGCCCGTCTGCATCGGTGAGTGCCCAGGGGTGA
TCCGCAATGCCACCACCGGCCGCATCGTCTCTCCAGGCTTCCCGGGCAACTACAGCAACAACCTCACCTGTCACTGG
CTGCTTGAGGCTCCTGAGGGCCAGCGGCTACACCTGCACTTTGAGAAGGTTTCCCTGGCAGAGGATGATGACAGGCT
CATCATTCGCAATGGGGACAACGTGGAGGCCCCACCAGTGTATGATTCCTATGAGGTGGAATACCTGCCCATTGAGG
GCCTGCTCAGCTCTGGCAAACACTTCTTTGTTGAGCTCAGTACTGACAGCAGCGGGGCAGCTGCAGGCATGGCCCTG
CGCTATGAGGCCTTCCAGCAGGGCCATTGCTATGAGCCCTTTGTCAAATACGGTAACTTCAGCAGCAGCACACCCAC
CTACCCTGTGGGTACCACTGTGGAGTTCAGCTGCGACCCTGGCTACACCCTGGAGCAGGGCTCCATCATCATCGAGT
GTGTTGACCCCCACGACCCCCAGTGGAATGAGACAGAGCCAGCCTGCCGAGCCCGTGTGCAGCGGGGAGATCACAGAC
TCGGCTGGCGTGGTACTCTCTCCCAACTGGCCAGAGCCCTACGGTCGTGGGCAGGATTGTATCTGGGGTGTGCATGT
GGAAGAGGACAACAGCCATCATGCTGGACATCCGAGTGCTGCGCATAGGCCCTGGTGATGTGCTTACCTTCTATGATG
GGGATGACCTGACGGCCCGGGTTCTGGGCCAGTACTCAGGGCCCCGTAGCCACTTCAAGCTCTTTACCTTACCTGGCT
GATGTCACCATTCAGTTCCAGTCGGACCCCGGGACCTCAGTGCTGGGCTACCAGCAGGGCTTCGTCATCCCACTTCTT
TGAGGTGCCCCGCAATGACACATGTCCGGAGCTGCCTGAGATCCCCAATGGCTGGAAGAGCCCATCGCAGCCTGAGC
TAGTGCACGGCACCGTGGTCACTTACCAGTGCTACCCTGGCTACCAGGTAGTGGGATCCAGTGTCCTCATGTGCCAG
TGGGACCTAACTTGGAGTGAGGACCTGCCCTCATGCCAGAGGGTGACTTCCTGCCACGATCCTGGAGATGTGGAGCA
CAGCCGACGCCTCATATCCAGCCCCAAGTTTCCCGTGGGGGCCACCGTGCAATATATCTGTGACCAGGGTTTTGTGC
TGATGGGCAGCTCCATCCTCACCTGCCATGATCGCCAGGCTGGCAGCCCCAAGTGGAGTGACCGGGCCCCTAAATGT
CTCCTGGAACAGCTCAAGCCATGCCATGGTCTCAGTGCCCCTGAGGATTGGTGCCCGAAGTCCTGAGAAGCAGCTACA
CCCAGCAGGGGCCACCATCCACTTCTCGTGTGCCCCTGGCTATGTGCTGAAGGGCCAGGCCAGCATCAAGTGTGTGC
CTGGGCACCCCTCGCATTGGAGTGACCCCCCACCCATCTGTAGGGCTGCCTCTCTGGATGGTTCTACAACAGTCGCA
GCCTGGATGGTTGCCAAGGCACCTGCTGCCTCCAGCACCCTGGATGCTGCCCACATTGCAGCTGCCATCTTCTTGCC
ACTGGTGGCGATGGTGTTGTTGGTAGGAGGTGTATACTTCTACTTCTCCAGGCTCCAGGGAAAAAGCTCCCTGCAGC
TGCCCCGCCCCCGCCCCCGCCCCTACAACCGCATTACCATAGAGTCAGCGTTTGACAATCCAACTTACGAGACTGGA
TCTCTTTCCTTTGCAGGAGACGAGAGAATATGAAGTCTCCATCTAGGTGGGGCAGTCTAGGGAAGTCAACTCAGAC
TTGCACCACAGTCCAGCAGCAAGGCTCCTTGCTTCCTGCTGTCCCTCCACCTCCTGTATATACCACCTAGGAGGAGA
TGCCACCAAGCCCTCAAGAAGTTGTGCCCTTCCCCGCCTGCGAGCCCACCATGGCCTATTTTCTTGGTGTCATTGC
CCACTTGGGGCCCTTCATTGGGCCCCATGTCAGGGGGCATCTACCTGTGGGAAGAACATAGCTGGAGCACAAGCATCA
ACAGCCGGCATCCTGAGCCTCCTCATGCCCCTGGACCAGCCTGGAACACACTAGCAGAGCAGGAGTACCTTTCTCCAC
ATGACCACCATCCCGCCCTGGCATGGCAACCTGCAGCAGGATTAACTTGACCATGGTGGGAACTGCACCCAGGGTACT
CCTCACAGCGCCATCACCAATGGCCAAAACTCCTCTCAACGGTGACCTCTGGGTAGTCCTGGCATGCCAACATCAGC
CTCTTGGGAGGTCTCTAGTTCTCTAAAGTTCTGGACAGTTCTGCCTCCTGCCCTGTCCCAGTGGAGGCAGTAATTCT
AGGAGATCCTAAGGGGTTCAGGGGGACCCTACCCCCACCTCAGGTTGGGCTTCCCTGGGCACTCAGCTCCACACCA
AAGCAGGACACGCCATTTTCCACTGACCACCCTATACCCTGAGGAAAGGGAGACTTTCCTCCGATGTTT
```

FIGURE 74B

```
ATTTAGCTGTTGCAAACATCTTCACCCTAATAGTCCCTCCTCCAATTCCAGCCACTTGTCAGGCTCTCCTCTTGACC
ACTGTGTTATGGGATAAGGGGAGGGGGTGGGCATATTCTGGAGAGGAGCAGAGGTCCAAGGACCCAGGAATTTGGCA
TGGAACAGGTGGTAGGAGAGCCCCAGGGAGACGCCCAGGAGCTGGCTGAAAGCCACTTTGTACATGTAATGTATTAT
ATGGGGTCTGGGCTCCAGCCAGAGAACAATCTTTTATTTCTGTTGTTTCCTTATTAAAATGGTGTTTTTGGAAAAAAA
```

FIGURE 75A

```
GCTGAGCCTGAGCCCGACCCGGGGCGCCTCCCGCCAGGCACCATGGTGCAGAAGTCGCGCAACGGCGGCGTATACCC
CGGCCCGAGCGGGGAGAAGAAGCTGAAGGTGGGCTTCGTGGGGCTGGACCCCGGCGCGCCCGACTCCACCCGGGACG
GGGCGCTGCTGATCGCCGGCTCCGAGGCCCCCAAGCGCGGCAGCATCCTCAGCAAACCTCGCGCGGGCGGCGCGGGC
GCCGGGAAGCCCCCCAAGCGCAACGCCTTCTACCGCAAGCTGCAGAATTTCCTCTACAACGTGCTGGAGCGGCCGCG
CGGCTGGGCGTTCATCTACCACGCCTACGTGTTCCTCCTGGTTTTCTCCTGCCTCGTGCTGCTGTGTTTTCCACCA
TCAAGGAGTATGAGAAGAGCTCGGAGGGGGCCCTCTACATCCTGGAAATCGTGACTATCGTGGTGTTTGGCGTGGAG
TACTTCGTGCGGATCTGGGCCGCAGGCTGCTGCTGCCGGTACCGTGGCTGGAGGGGGCGGCTCAAGTTTGCCCGGAA
ACCGTTCTGTGTGATTGACATCATGGTGCTCATCGCCTCCATTGCGGTGCTGGCCGCCGGCTCCCAGGGCAACGTCT
TTGCCACATCTGCGCTCCGGAGCCTGCGCTTCCTGCAGATTCTGCGGATGATCCGCATGGACCGGCGGGGAGGCACC
TGGAAGCTGCTGGGCTCTGTGGTCTATGCCCACAGCAAGGAGCTGGTCACTGCCTGGCTGGATCCGGCTTCCTTTGTCT
CATCCTGGCCTCGTTCCTGGTGTACTTGGCAGAGAAGGGGGAGAACGACCACTTTGACACCTACGCGGATGCACTCT
GGTGGGGCCTGATCACGCTGACCACCATTGGCTACGGGGACAAGTACCCCCAGACCTGGAACGGCAGGCTCCTTGCG
GCAACCTTCACCCTCATCGGTGTCTCCTTCTTCGCGCTGCCTGCAGGCATCTTGGGGTCTGGGTTTGCCCTGAAGGT
TCAGGAGCAGCACAGGCAGAAGCACTTTGAGAAGAGGCGGAACCCGGCAGCAGGCCTGATCCAGTCGGCCTGGAGAT
TCTACGCCACCAACCTCTCGCGCACAGACCTGCACTCCACGTGGCAGTACTACGAGCGAACGGTCACCGTGCCCATG
TACAGACTTATCCCCCCCGCTGAACCAGCTGGAGCTGCTGAGGAACCTCAAGAGTAAATCTGGACTCGCTTTCAGGAA
GGACCCCCCGCCGGAGCCGTCTCCAAGCCAGAAGGTCAGTTTGAAAGATCGTGTCTTCTCCAGCCCCCGAGGCGTGG
CTGCCAAGGGGAAGGGGTCCCCGCAGGCCCAGACTGTGAGGCGGTCACCCAGCGCCGACCAGAGCCTCGAGGACAGC
CCCAGCAAGGTGCCCAAGAGCTGGAGCTTCGGGGACCGCAGCCGGGCACGCCAGGCTTTCCGCATCAAGGGTGCCGC
GTCACGGCAGAACTCAGAAGAACGAAGCCTCCCCGGAGAGGACATTGTGGATGACAAGAGCTGCCCCTGCGAGTTTG
TGACCGAGGACCTGACCCCGGGCCTCAAAGTCAGCATCAGAGCCGTGTGTGTCATGCGGTTCCTGGTGTCCAAGCGG
AAGTTCAAGGAGAGCCTGCGGCCCTACGACGTGATGGACGTCATCGAGCAGTACTCAGCCGGCCACCTGGACATGCT
GTCCCGAATTAAGAGCCTGCAGTCCAGAGTGGACCAGATCGTGGGGCGGGGCCCAGCGATCACGGACAAGGACCGCA
CCAAGGGCCCGGCCGAGGCGGAGCTGCCCGAGGACCCCAGCATGATGGGACGGCTCGGGAAGGTGGAGAAGCAGGTC
TTGTCCATGGAGAAGAAGCTGGACTTCCTGGTGAATATCTACATGCAGCGGATGGGCATCCCCCCGACAGAGACCGA
GGCCTACTTTGGGGCCAAAGAGCCGGAGCCGGCGCCGCCGTACCACAGCCCGGAAGACAGCCGGGAGCATGTCGACA
GGCACGGCTGCATTGTCAAGATCGTGCGCTCCACGGGCCAGAAGAACTTCTCGGCGCCCCGGCGCGCG
CCCCCTGTCCAGTGTCCGCCCTCCACCTCCTGGCAGCCACAGAGCCACCCGCGCCAGGGCCACGGCACCTCCCCCGT
GGGGGACCACGGCTCCCTGGTGCGCATCCCGCCGCCCTGCCCACGAGCGGTCGCTGTCCGCCTACGGCGGGGGCA
ACCGCGCCAGCATGGAGTTCCTGCGGGCAGGAGGACACCCCGGGCTGCAGGCCCCCGAGGGGACCCTGCGGGACAGC
GACACGTCCATCTCCATCCCGTCGTGGACCACGAGGAGCTGGAGCGTTCCTTCAGCGGCTTCAGCATCTCCCAGTC
CAAGGAGAACCTGGATGCTCTCAACAGCTGCTACGCGGCCGTGGCGCCTTGTGCCAAAGTCAGGCCCTACATTGCGG
AGGGAGAGTCAGACACCGACTCCGACCTCTGTACCCGTGCGGGCCCCCGCCACGCTCGGCCACCGGCGAGGGTCCC
TTTGGTGACGTGGGCTGGGCCGGGCCCAGGAAGTGAGGCGGCGCTGGGCCAGTGGACCCGCCCGCGGCCCTCCTCAG
CACGGTGCCTCCGAGGTTTTGAGGCGGGAACCCTCTGGGGCCCTTTTCTTACAGTAACTGAGTGTGGCGGGAAGGGT
GGGCCCTGGAGGGGCCCATGTGGGCTGAAGGATGGGGGCTCCTGGCAGTGACCTTTTTACAAAAGTTATTTTCCAACA
GGGCACTCCCAGGCCCTGTCGCCATTGAGGTGCCTCCGCTGGGCTGTCTCCTCACCCCTCCCTGTGCTGGAGCCTGT
CCCAAAAAGGTGCCAACTGGGGAGGCCTCGGAAGCCACTGTCCAGGCTCCCACTGCCTGTCTGCTCTGTTCCCAAAGG
CAGCGTGTGGGCCTCGGGCCCTGCGGTGGCATGAAGCATCCCTTCTGGTGTGGGCATCGCTACGTGTTTTGGGGGC
AGCGTTTCACGGCGGTGCCCTTGCTGTCTCCCTTGGGCTGGCTCGAGCCTGGGGTCCATGTCCCTTTGCCGTCCCGT
CATGGGGCAGGGAATCCATAGCGGGGCCCACAGGCAGGGGTATGAGTGCGTCCCACCCAACGCAGCACCAGCCCCGG
CCACCGCTCCCCGTGTCCCCAGTTCCGTCTCAGCTACCTGGACTCCAGGACCCTGGAGAAGGGAGACCTGGCAGTGG
AGGGAGGCTGTGCTGTGTCGTGCCCCCTGCAGGTGTGACCCCGGCCTGCTCTTTCCTCCCCCGCCCAGGTGGGCCCGCC
TGCTCTTTCCTCCCCCCACCAGTATGGCCCCCACCTGCTCTTTCCTXCCCCCCCAAGGTGTGGCCCCACCTGTTCTTTC
CTCCCCTGCCGAGGTGTGACCCCACCTGCTCTTTCCTCCCTCCCAGTATGGCCCCACCTGCTCTTTCCTCCCCCGAG
GTGAGGCCCCGCCTGCTCTTTCCTCCCATGGGAGCCGCTGAGGCGTGCGCACCTGGGCACAGGTTGGGGCTCTGCAG
GATGAGGAAGACAGGCCAATCCCTTCCCTCCCAGAAGCTGGCCGCCCAGCAGGAGGGACTGAGGCCCAGACTCATGTC
CAGCAAGGAACGTGTGGTGTGTCCCCTGGGAAGTCTCTGGGCCCTGGGGAAGAGGGAAGGTGCACGTCCTGGGATGGT
TGCGGGGCCCTGTTTTGGGAGACAAAGGGGTAGAGGGTCTGTCTTGXGCCCCCCCAGACTCTAGCCCGAGCAGTGCA
GCCACCTACTGCCCCACCTCAGAGAAGTGCAGCGGGAAGGAGGCTGGAGGTGGTGCGCGCCTGCCTCGGGTGTCTGC
GTGAATGAGCGTGGCCAAGGACCAGTGCCACCTCATGGCAAAGAGCTCCCGCAGTGTTTGTTAGAGTGCACATCCTA
CGTGCCCACTGGCACACACACGTGCTCACATACAXGTCXGXGTACAGGXGTACACATGCAXGCTTGCAC
```

FIGURE 75B

```
ACATGCACACAGACCACATAGCACACATGTGCACTGACCACACCTGTATAGACCATGCACAGTACACATACGTGCAT
ACACATGCCTGCATACAGGCATACACATGCACGCTTACATGTACACGTGCACAGATCACACACATGCACACACGTGT
AGCTCACACACAGTATACACATACACAAGTGCACAGACCACACACAGCACTAACACATGCACACACAAAGTGCATAG
GCCACACAGCACATGCACACAGGTGCACAGACCACACAGCACACACAAGTGCACAGAGCACACTGCACACATGCACA
CACACACGCGTGCATGCACACTCCTCGCACTTCCAGCCTTGGAGCCCTTCTGTCTCTGGTCTTTCTCTTTGACCCTG
CTGAGTGTAAGCTGCCTGGGGAGGGGCTACAAGGAGTAATTGTGGCTTTAGGGGTCGTGGTGATGCTGGAATGTCAA
GCGCCGTCGTGGGGTATCCGACTGTCCGGGCTCCTGGTCCGCAGTGGCAGAGCGCCAGGCAGAGCCAATCAGGGTCT
CGTGCTGCCCTTCCCCCCCACAGCCTGGCAGCCATCCAGAGGAGGGGCTCTACCAGATGCCAAGGTGCCCCGGTGTC
TGTATGGGTGTCCGGTTGGGTCCTGTGTTTGGTCTGCCCTGGAGGTGGCTGGGCCCTCCTGGGATGGGTGGCTCAGC
CTCGAATCCCAGGCCCCAGCCCAGGCAGGTGCTGCTGCCTGTTGTGGTTTCCTGGCCCAGCTTCTCCTTCTCCCTCT
GCATAAAATCACAGTCCGTGAGTCTTCCAGCTGCCACCACGGCTGGGACACGCTGGGGGAGGGCTCCTCCCATGCCT
CCTGCACACAGCCGTCTGAGCAGGGCAGGTGCCAACACCCCCCACCGGAGACACGCTGCCCCTCAGCGATGCCCCTA
CCTTTTGGGGGGCCTCGTCTCAAGCCCCCCCTTGGAGGCTGAAATCACCCCAGGCACTGTGAGGGCTTCTCCAGGGG
GACACCCTTTGAGCTGTGGGTCTGATCACCCCAAGTCCCGCACACGGAGGAGAGGCACAGCCAGGGCGTGTGGTTTA
ATGTTTGCCCCTTCGGGGCTGGAGGTCTCAGTGTTTCTAGATTCCAGACCCTGCTGCCAGAGAGACCTGCTGCCGGA
GAGAAGGGGAGGAGGACTCCAGCTGGGCTCGGTCCCCCACAGTCAGGGACCCCCATAAAGGACACCCCCTTCTCTCT
AGAAAGAGCTGGGCTCTCAGCTATTTCTAGTTGCTTCCCAGAAGCCGAGGAGCAGAAGGAGCTGTGAGAGCTTTGCA
GAAACGCCCTTGTCCCCGCCCTCCTGAGCTATGAATGCCGTACAGAGCAGAGGCTGGGGCATTGGCAAGATCACAGG
TTGATGCTGCACAGCCCCATTGACACAAACCCTCAAAGCAGACGTGAGAGGGACGGTTCACAAAGCTTGGACCTGCC
GTGGAGGGTGCCCGGCAGACGTGGCGTGAGAGGGACGGCTCACGAGGCTTGGACCTGCTGTGGAGGGTGCCCAGCAG
ACGTGGTGTGAGAGGAACGGCTCACGAGACTTGGACCTGGTGGAGGGTGCCCAGCAGACGTGGTGTGAGAGGGACGG
CTCACAGGGCTTGGACCGGAGAGAGATGGCTCATGAGACTTGGACCTGCCGTGGAGGGTGCCCAGCAGACGTGGTAT
GAGAGGGATGGCTCACGAGGCTTGGACCTGGTGGAGGGTGCCCGGCAGACGTGTGAGAGGGACGGTTCACAAGGCTT
GGACCTGCCATGGAGGGTGCCCAGCAGACGTGGTGTGAGAGGGACAGCTCACGAGGCTTGGACCTGCCGTGGAGGGT
GCCCAGCAGGGGGCTGAGCTCTGAGGGGTGGGTGCTCAGTGCACGGGTGCCCCCAGTGTCCTCTGATCCTGTCCGGT
GCCTCCCCCAACCCCCACACCCATGCAGAACTCCCAGGTCACATGCACGTATGTCCAGGGCATGGGGGTGGCGTGAA
GAGGCCTGGTCAGGGCCTTTAGGGGCTGCAGGACGGAATGGCCACCTGGGGAGCCTGTGTGGCTGTGCCGGGCAGCC
ATCCTGCATTCCCACCCAGCGCGCAGTCTCCACCTCGGCCCCAGCAAAGCGCTAAGCAGCCGGAGAGACAGCCAGGG
CGGCTTCCTGAAGGATGTGGGATGGTGGACTCCGGGGTCGAGGGAATACGCAGGTTCCTGTCCTCCGGGAGACCTAG
AGAAGCTGCACACCCAGGAGCTTTCCATGACCCGGGAGCATGAGTGAATGGGGGGTTCCAGTTTGCTGAACTTTGCT
GTCTTGTAAGGGTGGGGGCTGACGGCCGACCCTGGGAGGAGGTGACACCGCAGGGGGAGGTTGTGGGCAACGGTGGA
GGAGGAGAGACGGGAGGGGACCATTTGGGATGGAGGGGCCTCTTCAGAGTTTTAAAAGGCGTTTGTGGGGTGGAGTT
GAGTGTGCTCTGGGCTTGGACACTTGCCGTGGTGCCCCTGGCTGGCCGAGGAGACTGGCTCTGGCCAGGGCCCCGTC
CTGAGAGGTCCTCAGCGTCTGACTCTCGGCCAGGCGCCAGCAAGGAGGGGCCGGTCCCCGGGGCTACCAGGCAGGCA
CGTGCACATCGCCATCGCCACACGCCAACTCCGCCTGGGTTTTACAAAGTCGTTGCCTTAATGCATGTGGACAGGAA
CTCCCTGAGGTCGCCCCATGCCCCCTGGCTGTGCCAGGTACGGACGCCCTGGACCCTGCGAACAGGTGGGGCGGGCG
AGGGGCCCAAGGACGCAGGGCACGCGCAGGACACGCGGAGGGGGTCTCACGGAGGGGTCTCGCACTGAGGCGCC
CAGAGCTGGTGGTCCCGCTGGACGCCATCCCTCTGCCCGGGATCCACACGGCCCACGTGTGCCCGCCATGCCCGCGC
CCCACGCCATTGCAGTCTTCCATCCTCTGGCCGTGACGGTGGCTGCAGCTTCCCCATTTGCGCCCGTTGCCTCTGGCT
GTCTGCACTTTTGTTCATGCTCCAAAGAACATTTCATAATGCCTTCAGTACCGACGTACACTTCTGACCATTTTGTA
TGTGTCCTTGTGCCGTAGTGACCAGGCCTTTTTTTGGTGGATGTGTTACCCCGCACACTTCAATCTCAACTTTGTGC
ACCGTCCATTTTCTAGGGATAGACGCCCAGGGAATGAACTCTAGTTTTCTAACAGATTAGCTGAGATATTAACTTAC
TCACACGGACAGGTTGATGCCAGAGCCGTAAGAATGCGCCAGTGCGGGTTTGCGGGGGACTTCGGGTGTGGGGTCCT
GCGGCCGCGATGGCCGTGGAAGGTTCTGGGGATCCCTGCTGCCACGGGGACGAGTTCGGACGCCAGGTGGACCTGTG
CACTCAGTAAAACGCAGTGATTCAAAAAAAAAAAAAA
```

FIGURE 76

```
GGCACGAGGCAAGTCCGGATGAAGAATTAAGAGAAAAAAAGTGAATATGGTTTTTGCTCACAGAATGGATAACAGCA
AGCCACATTTGATTATTCCTACACTTCTGGTGCCCCTCCAAAACCGCAGCTGCACTGAAACAGCCACACCTCTGCCA
AGCCAATACCTGATGGAATTAAGTGAGGAGCACAGTTGGATGAGCAACCAAACAGACCTTCACTATGTGCTGAAACC
CGGGGAAGTGGCCACAGCCAGCATCTTCTTTGGGATTCTGTGGTTGTTTTCTATCTTCGGCAATTCCCTGGTTTGTT
TGGTCATCCATAGGAGTAGGAGGACTCAGTCTACCACCAACTACTTTGTGGTCTCCATGGCATGTGCTGACCTTCTC
ATCAGCGTTGCCAGCACGCCCTTTCGTCCTGCTCCAGTTCACCACTGGAAGGTGGACGCTGGGTAGTGCAACGTGCAA
GGTTGTGCGATATTTTCAATATCTCACTCCAGGTGTCCAGATCTACGTTCTCCTCTCCATCTGCATAGACCGGTTCT
ACACCATCGTCTATCCTCTGAGCTTCAAGGTGTCCAGAGAAAAAGCCAAGAAAATGATTGCGGCATCGTGGATCTTT
GATGCAGGCTTTGTGACCCCTGTGCTCTTTTTCTATGGCTCCAACTGGGACAGTCATTGTAACTATTTCCTCCCCTC
CTCTTGGGAAGGCACTGCCTACACTGTCATCCACTTCTTGGTGGGCTTTGTGATTCCATCTGTCCTCATAATTTTAT
TTTACCAAAAGGTCATAAAATATATTTGGAGAATAGGCACAGATGGCCGAACGGTGAGGAGGACAATGAACATTGTC
CCTCGGACAAAAGTGAAAACTATCAAGATGTTCCTCATTTTAAATCTGTTGTTTTTGCTCTCCTGGCTGCCTTTTCA
TGTAGCTCAGCTATGGCACCCCCATGAACAAGACTATAAGAAAAGTTCCCTTGTTTTCACAGCTATCACATGGATAT
CCTTTAGTTCTTCAGCCTCTAAACCTACTCTGTATTCAATTTATAATGCCAATTTTCGGAGAGGGATGAAAGAGACT
TTTTGCATGTCCTCTATGAAATGTTACCGAAGCAATGCCTATACTATTACAACAAGTTCAAGGATGGCCAAAAAAAA
CTACGTTGGCATTTCAGAAATCCCTTCCATGGCCAAAACTATTACCAAAGACTCGATCTATGACTCATTTGACAGAG
AAGCCAAGGAAAAAAAGCTTGCTTGGCCCATTAACTCAAATCCACCAAATACTTTTGTCTAAGTTCTCATTCTTTCA
ATTGTTATGCACCAGAGATTAAAAAGCTTTAACTATAAAAACAGAAGCTATTTACATATTTGTTTTCACTCAACTTT
CCAAGGGAAATGTTTTATTTTGTAAAATGCATTCATTTGTTTACTGT
```

FIGURE 77

GCC**ATG**GCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGC
CCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCG
AGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAG
TGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGGCAGCACAGACGGGAGCAGTGGGCTGCCCACCAT
GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGG
CATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAG
GAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGC
CGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACA
GCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGC
CGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCT
GGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTG
ACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTG
GGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGG
GCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAG
ACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG
CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACT
GCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGC
CCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATC
AGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGA
CGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCA
CCCGGCCAGTCATCCTGGCTATGGACACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAG
TCGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGA
GGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCG
CCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTG
CATGACACCCACGATGCACTCAAGGAGGTGGTGCACGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG
GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTA
AGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTC
ACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCA
GAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGG
TCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA
CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTC
CAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGG
CTGCAGCAGTGGCTGGACTGGACTGGGCCAGCAGCACGGACCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC**TGA**GGCCGGCCA
GGCCCGACACCTACACTCTCACCAGCTTTGGCACCCACCAAGGACAGGCAGGAGACCGGACAGGGGCCCTTCCCCACA
CCGGGGAGAGCTGCTCGGACAGGCCCCTCCCGGCCGAAGCTGTCCCTTAATGCTGGTCCTTCAGACCCTGCCCCGAA
CTCCCACCTCTCCATGGCCTGCCTAGCCAGGCTGGCACTGCCACTCACACTCGGCCCCAGGGCCCAGGAGGGACAGT
GCCTGGAGCCTGGGCCAGGCCCAGCCCATCTGTGTGTGTGTATGTGCGTGTGATGCTACCTCTCCTCCCGTCCCTCT
CCAGGGGCCCCGCATACACACGGCCATGCACGCACACACTGGGCCTGGGCCAGGGCCCCAGAGCTCCTGCCTGAGCT
GGACCTTATGCAAACATTTCTGTGCCTGCTGGGTAGGGGCACGTCTGAGGGGCCCTGCTCCAAGCCTGCAGGACCGA
GGGCCACAGCCGGACAGGGGGTAGCCCCTGGATTCAGGCACACGACCACCACACGAGCACGTGCCACGCATGCCTCG
TGTGCTCATCTCACACACACCCCCCTCCCGGGTCACGCAGACACCCCCCAACCACACATCTCATGCCGTACACCT
GAGGCTGCTCACGTCTCACGCCCAGTGTTGGTGCACATTTGCCTCTCACATGCTGCCCTCTCCACCCACCCAGGGAC
ACCCCACGGCTCCTCCCTGCCCCTCCCCCAGCCTTGAGGTGCCCTGCCCGGCGGGGCCTGTGAATATGCAA
TGGGAGTCCCAGGCTGTACAGTGGTGAGTGTGTGTGGCGTGGCCTGCCCGTCCCCCAGGGCTGGCTGGTGCCCCAC
GCGGGGCCTGTCATGTGAAGCTCGTGTCCTGACTTTGTCTTAAGTGCATTCACGCACTTACTCTTGGCCTTATGTAC
ACAGCCTTGCCCGGCCGCCGGGGCCACATAGGGGTTTTATCGGGCGTGAATGTAAATAAATTATATATATATATTGCT
AAAAAAAAAAAAAAAAAAA

FIGURE 78A

```
TAAGGTAATTCGTATGCAAGAAGCTACACGTAATTAAATGTGCAGGATGAAAAGATGGCACAGGCACTGTTGGTACC
CCCAGGACCTGAAAGCTTCCGCCTTTTTACTAGAGAATCTCTTGCTGCTATCGAAAAACGTGCTGCAGAAGAGAAAG
CCAAGAAGCCCAAAAAGGAACAAGATAATGATGATGAGAACAAACCAAAGCCAAATAGTGACTTGGAAGCTGGAAAG
AACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACCTGGATCCCTACTATAT
CAATAAGAAAACTTTTATAGTAATGAATAAAGGAAAGGCAATTTTCCGATTCAGTGCCACCTCTGCCTTGTATATTT
TAACTCCACTAAACCCTGTTAGGAAAATTGCTATCAAGATTTTGGTACATTCTTTATTCAGCATGCTTATCATGTGC
ACTATTTTGACCAACTGTGTATTTATGACCTTGAGCAACCCTCCTGACTGGACAAAGAATGTAGAGTACACATTCAC
TGGAATCTATACCTTTGAGTCACTTATAAAAATCTTGGTAAGAGGGTTTTGCTTAGAAGATTTTACGTTTCTTCGTG
ATCCATGGAACTGGCTGGATTTCAGTGTCATTGTGATGGCGTATGTAACAGAATTTGTAAGCCTAGGCAATGTTTCA
GCCCTTCGAACTTTCAGAGTCTTGAGAGCTCTGAAAACTATTTCTGTAATTCCAGGTTTAAAGACCATTGTGGGGGC
CCTGATCCAGTCGGTAAAGAAGCTTTCTGATGTGATGATCCTGACTGTGTTCTGTCTGAGCGTGTTTGCTCTCATTG
GGCTGCAGCTGTTCATGGGCAATCTGAGGAATAAATGTTTGCAGTGGCCCCCAAGCGATTCTGCTTTTGAAACCAAC
ACCACTTCCTACTTTAATGGCACAATGGATTCAAATGGGACATTTGTTAATGTAACAATGAGCACATTTAACTGGAA
GGATAACATTGGAGATGACAGTCACTTTTATGTTTTGGATGGGCAAAAAGACCCTTTACTCTGTGGAAATGGTTCAG
ATGCAGGCCAGTGTCCAGAAGGATACATCTGTGTGAAGGCTGGTCGAAACCCCAACTATGGCTACACAAGCTTTGAC
ACCTTTAGCTGGGCTTTCCTGTCTCTATTTCGACTCATGACTCAAGACTACTGGGAAAATCTTTACCAGTTGACATT
ACGTGCTGCTGGGAAAACATACACGATATTTTTTGTCCTGGTCATTTTCTTGGGCTCATTTTATTTGGTGAATTTGA
TCCTGGCTGTGGTGGCCATGGCCTATGAGGAGCAGAATCAGGCCACCTTGGAAGAAGCAGAACAAAAAGAGGCCGAA
TTTCAGCAGATGCTCGAACAGCTTAAAAAGCAACAGGAAGAAGCTCAGGCAGTTGCGGCAGCATCAGCTGCTTCAAG
AGATTTCAGTGGAATAGGTGGGTTAGGAGAGCTGTTGGAAAGTTCTTCAGAAGCATCAAAGTTGAGTTCCAAAAGTG
CTAAAGAATGGAGGAACCGAAGGAAGAAAAGAAGACAGAGAGACAGCACCTTGAAGGAAACAACAAAGGAGAGAGAGAC
AGCTTTCCCAAATCCGAATCTGAAGACAGCGTCAAAAGAAGCAGCTTCCTTTTCTCCATGGATGGAAACAGACTGAC
CAGTGACAAAAAATTCTGCTCCCCTCATCAGTCTCTCTTGAGTATCCGTGGCTCCCTGTTTTCCCCAAGACGCAATA
GCAAAACAAGCATTTTCAGTTTCAGAGGTCGGGCAAAGGATGTTGGATCTGAAAATGACTTTGCTGATGATGAACAC
AGCACATTTGAAGACAGCGAAACGAGGAGAGACTCACTGTTTGAGCCGCACAGACATGGAGAGCGACGCAACAGTAA
CGGCACCACCACAGAAACGGAGGTCAGAAAGAGAAGGTTAAGTTCTTACCAGATTTCAATGGAGATGCTGGAGGATT
CTTCTGGAAGGCAAAGAGCCGTGAGCATAGCCAGCATTCTGACCAACACAATGGAAGAACTTGAAGAATCTAGACAG
AAATGTCCGCCATGCTGGTATAGATTTGCCAATGTGTTCTTTGATCTGGGACTGCTGTGATGCATGGTTAAAAGTAAA
ACATCTTGTGAATTTAATTGTTATGGATCCATTTGTTGATCTTTGCATCTCACTATTTTGCATTGTCTTAAATACCCTCT
TTATGGCCATGGAGCACTACCCCATGACTGAGCAATTCAGTAGTGTGTTGACTGTAGGAAACCTGGTCTTTACTGGG
ATTTTCACAGCAGAAATGGTTCTCAAGATCATTGCCATGGATCCTCTTATTACTATTTCCAAGAAGGCTGGAATATCTT
TGATGGAATTATTGTCAGCCTCAGTTTAATGGAGCTTGGTCTGTCAAATGTGGAGGGATTGTCTGTACTGCGATCAT
TCAGACTGCTTAGAGTTTTCAAGTTGGCAAAATCCTGGCCCACACTAAATATGCTAATTAAGATCATTGGCAATTCT
GTGGGGGCTCTAGGAAACCTCACCTTGGTGTTGGCCATCATCGTCTTCATTTTTGCTGTGGTCGGCATGCAGCTCTT
TGGTAAGAGCTACAAAGAATGTGTCTGCAAGATCAATGATGACTGTACGCTCCCACGGTGGCACATGAAGGACGACTTCT
TCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGAGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTCGCT
GGCCAAACCATGTGCCTTATTGTTTTCATGTTGGTCATGGTCATTGGAAACCTTGTGGTTCTGAACCTCTTTCTGGC
CTTATTGTTGAGTTCATTTAGCTCAGACAACCTTGCTGCTACTGATGATGACAATGAAATGAATAATCTGCAGATTG
CAGTAGGAAGAATGCAAAAGGGAATTGATTATGTGAAAAATAAGATGCGGGAGTGTTTTCCAAAAAGCCTTTTTTAGA
AAGCCAAAAGTTATAGAAATCCATGAAGGCAATAAGATAGACAGCTGCATGTCCAATAATACTGGAATTGAAATAAG
CAAAGCGCTTAATTATCTTAGAGATGGGAATGGAACCACCAGTGGTGTAGGTACTGGAAGCAGTGTTGAAAAATACG
TAATCGATGAAAATGATTATATGTCATTCATAAACAACCCCAGCCTCACCGTCACAGTGCCAATTGCTGTTGGAGAG
TCTGACTTTGAAAACTTAAATACTGAAGAGTTCAGCAGTGAACTAGAAGCAGTAATAGAAGAGAAAGAGAAATTAAATGC
AACCAGCTCATCTGAAGGAAGCACAGTTGATGTTGTTCTACCCCGAGAAGGTGAACAAGCTGAAACTGAACCCCGAAAG
AAGACCTTAAACCGGAAGCTTGTTTTACTGAAGGATGTATTAAAAAGTTTCCATTCTGTCAAGTAAGTACAGAAGAA
GGCAAAGGGAAGATCTGGTGGAATCTTTCGAAAAACCTGCTACAGTATTGTTGAGCACAACTGGTTTGAGACTTTCAT
TGTGTTCATGATCCTTCTCAGTAGTGGTGCATTGGCCTTTGAAGATATATATACATTGAACAGCGAAGACGATATCAAAA
CCATGCTAGAATATGCTGACAAAGTCTTTACCTATATATTCATTCTTGGAAATGCTTCTCAAATGGGTTGCTTATGGA
TTTCAAACATATTTCACTAATGCCTGGTGCCGGCTAGATTTCTTTGATCGTTGATGTTTCTTTGGTTAGCCTGGTAGC
CAATGCTCTTGGCTACTCAGAACTCGGTGCCATCAAATCATTTACGGACATTAAGAGCTTTAAGACCTCTAAGAGCCT
TATCCCGGTTTGAAGGCATGAGGGTGGTTGTGAATGCTCTTGTTGGAGCAATTCCCTCTATCATGAATGTGCTGTTG
GTCTGTCTCATCTTCTGGTTGATCTTTAGCATCATGGGTGTGAATTTGTTTGCTGGCAAGTTCTACCAC
```

FIGURE 78B

TGTGTTAACATGACAACGGGTAACATGTTTGACATTAGTGATGTTAACAATTTGAGTGACTGTCAGGCTCTTGGCAA
GCAAGCTCGGTGGAAAAACGTGAAAGTAAACTTTGATAATGTTGGCGCTGGCTATCTTGCACTGCTTCAAGTGGTAAGTG

FIGURE 79A

```
AGCGAAGCGGAGGCATAAGCAGAGAGGATTCTGGAAAGGTCTCTTTGTTTTCTTATCCACAGAGAAAGAAAGAAAAA
AAATTGTAACTAATTTGTAAACCTCTGTGGTCAAAAAAAAAAAAAAAAAAAGCTGAACAGCTGCCAGAGGAAGAC
ACGTTATACCCTAACCATCTTGGATGCTGGGCTTTGTTATGCTGTAATTCATAAGGCTCTGTTTTATCAGAGATTAT
GGAGCAAGAAAACTGAAGCCAAGCCACATCAAGGTTTGACAGGGATGAGATACCTGTCAAGGATTCATAGTAGAGTG
GCTTACTGGGAAAGGAGCAAAGAATCTCTTCTAGGGATATTGTAAGAATAAATGAGATAATTCACAGAAGGGACCTG
GAGCTTTTCCGGAAAAAGGTGCTGTGACTATCTAAGGTAATTCGTATGCAAGAAGCTACACGTAATTAAATGTGCAG
GATGAAAAGATGCACAGGCACTGTTGGTACCCCCAGGACCTGAAAGCTTCCGCCTTTTTACTAGAGAATCTCTTGC
TGCTATCGAAAAACGTGCTGCAGAAGAGAAAGCCAAGAAGCCCAAAAAGGAACAAGATAATGATGATGAGAACAAAC
CAAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCA
GAGCCCCTGGAGGACCTGGATCCCTACTATATCAATAAGAAAACTTTTATAGTAATGAATAAAGAGAAAGGCAATTTT
CCGATTCAGTGCCACCTCTGCCTTGTATATTTAACTCCACTAAACCCTGTTAGGAAAATTGCTATCAAGATTTTGG
TACATTCTTTATTCAGCATGCTTATCATGTGCACTATTTTGACCAACTGTGTATTTATGACCTTGAGCAACCCTCCT
GACTGGACAAAGAATGTAGAGTACACATTCACTGGAATCTATACCTTTGAGTCACTTATAAAAATCTTGGCAAGAGG
GTTTTGCTTAGAAGATTTTACGTTTCTTCGTGATCCATGGAACTGGCTGGATTTCAGTGTCATTGTGATGGCATATG
TGACAGAGTTTGTGGACCTGGGCAATGTCTCAGCGTTGAGAACATTCAGAGTTCTCCGAGCACTGAAAACAATTTCA
GTCATTCCCAGGTTTAAAGACCATTGTGGGGCCCTGATCCAGTCGGTAAAGAAGCTTTCTGATGTGATGATCCTGAC
TGTGTTCTGTCTGAGCGTGTTTGCTCTCATTGGGCTGCAGCTGTTCATGGGCAATCTGAGGAATAAATGTTTGCAGT
GGCCCCCAAGCGATTCTGCTTTTGAAACCAACACCACTTCCTACTTTAATGGCACAATGGATTCAAATGGGACATTT
GTTAATGTAACAATGAGCACATTTAACTGGAAGGATTACATTGGAGATGACAGTCACTTTTATGTTTTGGATGGGCA
AAAAGACCCTTTACTCTGTGGAAATGGCTCAGATGCAGGCCAGTGTCCAGAAGGATACATCTGTGTGAAGGCTGGTC
GAAACCCCAACTATGGCTACACAAGCTTTGACACCTTTAGCTGGGCTTTCCTGTCTCTATTTCGACTCATGACTCAA
GACTATTGGGAAAATCTTTACCAGTTGACATTACGTGCTGCTGGGAAAACATACATGATATTTTTGTCCTGGTCAT
TTTCTTGGGCTCATTTTATTTGGTGAATTTGATCCTGGCTGTGGTGGCCATGGCCATGAGGAGCAGAATCAGGCCA
CCTTGGAAGAAGCAGAACAAAAAGAGGCCGAATTTCAGCAGATGCTCGAACAGCTTAAAAAGCAACAGGAAGAAGCT
CAGGCAGTTGCGGCAGCATCAGCTGCTTCAAGAGATTTCAGTGGAGTAGGTGGGTTAGGAGAGCTGTTGGAAAGTTC
TTCAGAAGCATCAAAGTTGAGTTCCAAAGGTGCTAAAGAATGGAGGAACCGGAGGAAGAAAAGAAGACAGAGAGAGAGC
ACCTTGAAGGAAACAACAAAGGAGAGAGAGACAGCTTTCCCAAATCCGAATCTGAAGACAGCGTCAAAAGAAGCAGC
TTCCTTTTCTCCATGGATGGAAACAGACTGACCAGTGACAAAAAATTCTGCTCCCCTCATCAGTCTCTCTTGAGTAT
CCGTGGCTCCCTGTTTTCCCCAAGACGCAATAGCAAAACAAGCATTTTCAGTTTCAGAGGTCGGGCAAAGGATGTTG
GATCTGAAAATGACTTTGCTGATGATGAACACAGCACATTTGAAGACAGCGAAAGCAGGAGAGACTCACTGTTTGTG
CCGCACAGACATGGAGAGCGACGCAACAGTAACGGCACCACCACTGAAACCGGAAGTCAGAAAGAGAAGGTTAAGCTC
TTACCAGATTTCAATGGAGATGCTGGAGGATTCCTCTGGAAGGCAAAGAGCCGTGAGCATAGCCAGCATTCTGACCA
ACACAATGGAAGAACTTGAAGAATCTAGACAGAAATGTCCGCCATGCTGGTATAGATTTGCCAATGTGTTCTTGATC
TGGGACTGCTGTGATGCATGGTTAAAAGTAAAACATCTTGTGAATTTAATTGTTATGGATCCATTTGTTGATCTTGC
CATCACTATTTGCATTGTCTTAAATACCCTCTTTTATGCCCATGAACGACTACCCCATGACTGAGCAATTCAGTAGTG
TGTTGACTGTAGGAAAACCTGGTCTTTACTGGGATTTTCACAGCAGAAATGGTTCTCAAGATCATTGCCATGGATCCT
TATTACTATTTCCAAGAAGGCTGGAATATCTTTGATGGAATTATTGTCAGCCTCAGTTTAATGGAGCTTGGTCTGTC
AAATGTGGAGGGATTGTCTGTACTGCGATCATTCAGACTGCTTAGAGTTTTCAAGTTGGCAAAATCCTGGCCCACAC
TAAATATGCTAATTAAGATCATTGGCAATTCTGTGGGGGCTCTAGGAAACCTCACCTTGGTGTTGGCCATCATCGTC
TTCATTTTTGCTGTGGTCGGCATGCAGCTCATTTTGTAAGAGCTACAAAGAAATGTGATCGATGGAAGAGTGGA
TACGCTCCCACGGTGGCACATGAACGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGAGAGTGGA
TAGAGACCATGTGGGACTGTATGGAGGTCGCTGGCCAAACCATGTGCCTTATTGTTTTCATGTTGGTCATGGTCATT
GGAAACCTTGTGGTTCTGAACCTCTTTTCTGGCCTTATTATTGAGTTCATTTAGCTCAGACAACCTTGCTGCTACTGA
TGATGACAATGAAATGAATAATCTGCAGATTGCAGTAGGAAGAATGCAAAAGGGGAATTGATTATGTGAAAAATAAGA
TGCGGGAGTGTTTCCAAAAAGCCTTTTTTAGAAAGCCAAAAGGTTATAGAAATCCATGAAGGGCAATAAGATAGACAGC
TGCATGTCCAATAATACTGGAATTGAAATAAGCAAAGAGCTTAATTATCTTAGAGATGGGAATGGAACCACCAGTGG
TGTAGGTACTGGAAGCAGTGTTGAAAAATACGTAATCGATGAAAATGATTATATGTCATTCATTCAAAAACAACCCAGCC
TCACCGTCACAGTGCCAATTGCTGTTGGAGAGTCTGACTTTGAAAACTTAAATACTGAAGAGTTCAGCAGTGAGTCA
GAACTAGAAGAAAGCAAAGAGAAATTAAATGCAACCAGCTCATCTGAAGGAAGCACAGTTGATGTTGTTCTACCCCG
AGAAGGTGAACAAGCTGAAACTGAACCCCGAAGAAGAACTTTAAACCCGGAAGCTTGTTTTACTGAAGGGTGTATTAAAA
AGTTTCCATTCTGTCAAGTAAGTACAGAAGAAGGCAAAGGGAAGATCTGGTGGAATCTTCGAAAAACCTGCTACACAGT
ATTGTTGAGCACAACTGGTTTGAGACTTTCATTGTGTTCATGATCCTTCTCAGTAGTGGTGCATTGGCCTTTGAAGA
TATATACATTGAACAGCGAAAGACTATCAAAACCATGCTAGAATATGCTGACAAAGTCTTTACCTATATATTCATTC
TGGAAATGCTTCTCAAATGGGTTGCTTATGGATTTCAAACATATTTCACTAATGCCTGGTGCTGGCTAGATTTTCTTG
ATCGTTGATGTTTCTTTGGTTAGCCTGGTAGCCAATGCTCTTGGCTACTCAGAACTCGGTGCCATCAAATCATTACG
GACATTAAGAGCTTTAAGCACCTCTAAGAGCCTTATCCCGGTTTGAAGGCATGAGGGTGGTTGTGAATGCTCTTGTTG
GAGCAATTCCCTCTATCATGAATGTGCTGTTGGTCTGTCTCATCTTCTGGTTGATCTTTTAGCATCATGGGTGTGAAT
TTGTTTGCTGGCAAGTTCTACCACTGTGTTAACATGACAACGGGTAACATGTTGACATTAGTGATGTTAACAATTT
GAGTGACTGTCAGGCTCTTGGCAAGCAAGCTCGGTGGAAAAACGTGAAAGTAAACTTTGATAATGTTGGCGCTGGCT
ATCTTGCACTGCTTCAAGTGGCCACATTTAAAGGCTGGATGGATATTATGTATGCAGCTGTTGATTCACGAGATGTT
AAACTTCAGCCTGTATATGAAGAAAATCTGTACATGTATTTATACTTTGTCATCTTTATCATCTTTGGGTCATTCTT
CACTCTGAATCTATTCATTGGTGTCATCATAGATAACTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCT
TTATGACAGAGGAACAGAAAAAAATATTACAATGCAATGAAGAAACTTGGATCCAAGAAACCTCAGAAACCCCATACCT
```

FIGURE 79B

```
CGCCCAGCAAACAAATTCCAAGGAATGGTCTTTGATTTTGTAACCAGACAAGTCTTTGATATCAGCATCATGATCCT
CATCTGCCTCAACATGGTCACCATGATGGTGGAAACGGATGACCAGGGCAAATACATGACCCTAGTTTTGTCCCGGA
TCAACCTAGTGTTCATTGTTCTGTTCACTGGAGAATTTGTGCTGAAGCTCGTTTCCCTCAGACACTACTACTTCACT
ATAGGCTGGAACATCTTTGACTTTGTGGTGGTGATTCTCTCCATTGTAGGTATGTTTCTGGCTGAGATGATAGAAAA
GTATTCTGTGTCCCCTACCTTGTTCCGAGTGATCCGTCTTGCCAGGATTGGCCGAATCCTACGTCTGATCAAAGGAG
CAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTGCTCTTC
CTGGTCATGTTTATCTATGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAAAAGGAAGCTGGAATTGATGACAT
GTTCAACTTTGAGACCTTTGACAACAGCATGATCTGCTTGTTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGC
TAGCACCTATTCTTAATAGTGCACCACCCGACTGTGACCCTGACACAATTCACCCTGGCAGCTCAGTTAAGGGAGAC
CGTGGGGACCCATCTGTTGGGATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTA
CATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCCCTGAGTGAGGATGACTTTGAGA
TGTTCTATGAGGTTTGGGAAAAGTTTGATCCCGATGCGACCCAGTTTATAGAGTTCTCTAAACTCTCTGATTTTGCA
GCTGCCCTGGATCCTCCTCTTCTCATAGCAAAACCCAACAAAGTCCAGCTTATTGCCATGGATCTGCCCATGGTCAG
TGGTGACCGGATCCACTGTCTTGATATTTTATTTGCCTTTACAAAGCGTGTTTTGTGTGAGAGTGGAGAGATGGATG
CCCTTCGAATACAGATGGAAGACAGGTTTATGGCATCAAACCCCTCCAAAGTCTCTTATGAGCCTATTACAACCACT
TTGAAACGTAAACAAGAGGAGGTGTCTGCCGCTATCATTCAGCGTAATTTCAGCTGTTATCTTTAAAGCAAAGGTT
AAAAAATATATCAAGTAACTATAACAAAGAGGCAATTAAAGGGAGGATTGACTTACCTATAAAACAAGACATGATTA
TTGACAAACTAAATGGGAACTCCACTCCAGAAAAAACAGATGGGAGTTCCTCTACCACCTCTCCTCCTTCCTATGAT
AGTGTAACAAAACCAGACAAGGAAAAGTTTGAGAAAGACAAACCAGAAAAAGAAAGCAAAGGAAAAGAGGTCAGAGA
AAATCAAAAGTAAAAAGAAACAAAGAAATTATCTTTGTGATCAATTGTTTACAGCCTATGAAGGTAAAGTATATGTGT
CAACTGGACTTCAAGAGGAGGTCCATGCCAAACTGACTGTTTTAACAAATACTCATAGTCAGTGCCTATACAAGACA
GTGAAGTGACCTCTCTGTCACTGCAACTCTGTGAAGCAGGGTATCAACGTTGACAAGAGGTTGCTGTTTTTATTACC
AGCTGACACTGCTGAGGAGAAACCCAATGGCTACCTAGACTATAGGGATAGTTGTGCAAAGTGAACATTGTAACTAC
ACCAAACACCTTTAGTACAGTCCTTGCATCCATTCTATTTTTAACTTCCATATCTGCCATATTTTTACAAAATTTGT
TCTAGTGCATTTCCATGGTCCCCAATTCATAGTTTATTCATAATGCTATGTCACTATTTTTGTAAATGAGGTTTACG
TTGAAGAAACAGTATACAAGAACCCTGTCTCTCAAATGATCAGACAAAGGTGTTTTGGCAGAGAGATAAAATTTTTG
CTCAAAACCAGAAAAAGAAATTGTAATGGCTACAGTTTCAGTTACTTCCCATTTTCTAGATGGCTTTAATTTTGAAAGT
ATTTTAGTCTGTTATGTTTGTTTCTATCTGAACAGTTATGTGCCTGTAAAGTCTCCTCTAATATTTAAAGGATTATT
TTTATGCAAAGTATTCTGTTTCAGCAAGTGCAAATTTTATTCTAAGTTTCAGAGCTCTATATTTAATTTAGGTCAAA
TGCTTTCCAAAAAGTAATCTAATAAATCCATTCTAGAAAAATATATCTAAAGTATTGCTTTAGAATAGTTGTTCCAC
TTTCTGCTGCAGTATTGCTTTGCCATCTTCTGCTCTCAGCAAAGCTGATAGTCTATGTCAATTAAATACCCTATGTT
ATGTAAATAGTTATTTTATCCTGTGGTGCATGTTTGGGCAAATATATATAGCCTGATAAACAACTTCTATTAAAT
CAAATATGTACCACAGTGTATGTGTCTTTGCAAGCTTCCAACAGGATGTATCCTGTATCATTCATTAAACATAGT
TTAAAGGCTATCACTAATGCATGTTAATATTGCCTATGCTGCTCTATTTTACTCAATCCATTCTTCACAAGTCTTGG
TTAAAGAATGTCACATATTGGTGATAGAATGAATTCAACCTGCTCTGTCATTATGTCAAGCAGAATAATTTGAAGC
TATTTACAAACACCTTTACTTTTGCACTTTTAATTCAACATGAGTATCATATGGTATCTCTCTGGATTTCAAGGAAA
CACACTGGATACTGCCTACTGACAAAACCTATTCTTCATATTTTGCTAAAAATAGTCTCTAAAACTTGTTTAAATATA
AATAATGTAAAAATATAATCAACTTTATTTGTCAGCATTTTGTACATAAGAAAATTATTTTCAGGTTGATGACATCA
CAATTTATTTTACTTTATGCTTTTGCTTTGATTTTTAATCACAATTCCAAACTTTTGAATCCATAAGATTTTTCAA
TGGATAATTTCCTAAAATAAAAGTTAGATAATGGGTTTTATGGATTTCTTTGTTATAATATATTTTCTACCATTCCA
ATAGGAGATACATTGGTCAAACACTCAAACCTAGATCATTTTCTACCAACTATGGTTGCCTCAATATAACCTTTTAT
TCATAGATGTTTTTTTTTATTCAACTTTTGTAGTATTTACGTATGCAGACTAGTCTTATTTTTTTTAATTCCTGCTGC
ACTAAAGCTATTACAAATATAACATAGTGACTTTGTTCTTTTTAGCCATGACATGAACAAAGTGGCAATTACCT
AACATGATATAAATTTTTGTTTTTGCACAAACCAAAAGTTTAATGTTAATTCTTTTTACAAAACTATTTACTGTAG
TGTATTGAAGAACTGCATGCAGGGAATTGCTATTGCTAAAAAGAATGGTGAGCTACCTGCATTATTGAGCCAAAAGAA
TAAATTTCATTTTTTATTGCATTTCACTTATTGGGCTCTGGGGTTTTTGTTTTTGTTTTTGCTGTTGGCAGTTTA
AAATATATATAATTAATTAAAAACCTGTGCTTGATCTGACATTTGTATACATAAAAGTTTACATGAATTTTACAACAAA
CTAGTGCATGATTCACCAAGCAGTACTACAGAACAAAGGCAAATTAAAAGCAGCTTTGTGAACTTTTATGTGTGCAA
AGGATCAAGTTCACATGTTCCAACTTTCAGGTTTGATAATAATAGTAGTAACCACCTACAATAGCTTTCAATTTCAA
TTAACTCCCTTGGCTATAAGCATCTAAACTCATCTTCTTTCAATATAATTGATGCTATCTCCTAATTACTTGGTGGC
TAATAAATGTTACATTCTTTGTTACTTAAATGCATTATATAAACTCCTATGTATACATAAGGTATTAATGATATAGT
TATTGAGAATTTATATTAACTTTTTTTTCAAGAACCCTTGGATTTATGTGAGGTCAAAACCAAACTCTTATTCTCAG
TGGAAAACTCCAGTTGTAATGCATATTTTTAAAGACAATTTGGATCTAAATATGTATTTCATAATTCTCCCATAATA
AATTATATAAGGTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 80

AAAAAAATTAAAAAAAAAAAAAAAAAACAGAAAAAAAAACATAGTACATGCCAAGATATTATTATGACAATTACAAAT
ACAAATAAATTATGATCTTTGACCTCAGCATATTTATTAACTAAAAGGGAAGATAAAACAGGCACATAACTATAACA
GGGGCACCAGTC**ATG**GGCGCCGCAGCCGCTCAGGCGCCTCTCGGGCTGCCTGCGGCCTCCGCTCGCCTTCTGCTGCT
AGCGACGTCGGTGCTTCTGCTCTTCGCCTTCTCTCTGCCCGGGAGCCGCGCGTCCAACCAGCCCCCGGGTGGTGGCG
GCGGCACGGGCGGGGACTGTCCCGGCGGCAAAGGCAAGAGCATCAACTGCTCAGAATTAAATGTGAGGGAGTCTGAC
GTAAGAGTTTGTGATGAGTCATCATGTAAATATGGAGGAGTCTGTAAAGAAGATGGAGATGGTTTGAAATGTGCATG
CCAATTTCAGTGCCATACAAATTATATTCCTGTCTGTGGATCAAATGGGGACACTTATCAAAATGAATGCTTTCTCA
GAAGGGCTGCTTGTAAGCACCAGAAAGAGATAACAGTAATAGCAAGAGGACCATGCTACTCTGATAATGGATCTGGA
TCTGGAGAAGGAGAAGAGGAAGGGTCAGGGGCAGAAGTTCACAGAAAACACTCCAAGTGTGGACCCTGCAAATATAA
AGCTGAGTGTGATGAAGATGCAGAAAATGTTGGGTGTGTATGTAATATAGATTGCAGTGGATACAGTTTTAATCCTG
TGTGTGCTTCTGATGGGAGTTCCTATAACAATCCCTGTTTTGTTCGAGAAGCATCTTGTATAAAGCAAGAACAAATT
GATATAAGGCATCTTGGTCATTGCACAGATACAGATGACACTAGTTTGTTGGGAAAGAAAGATGATGGACTACAATA
TCGACCAGATGTGAAAGATGCTAGTGATCAAAGAGAAGATGTTTATATTGGAAACCACATGCCTTGCCCTGAAAACC
TCAATGGTTACTGCATCCATGGAAAATGTGAATTCATCTATCTACTCAGAAGGGCTTCTTGTAGATGTGAATCTGGC
TACACTGGACAGCACTGTGAAAAGACAGACTTTAGTATTCTCTATGTAGTGCCAAGTAGGCAAAAGCTCACTCATGT
TCTTATTGCAGCAATTATTGGAGCTGTACAGATTGCCATCATAGTAGCAATTGTAATGTGCATAACAAGAAAATGCC
CCAAAAACAATAGAGGACGTCGACAGAAGCAAAACCTAGGTCATTTTACTTCAGATACGTCATCCAGAATGGTT**TAA**
ACTGATGACTTTTATATGTACACTGACCATGTGTATGTACATTTATTATGTCTTTTTTTAAAGAATGGAAATATTTA
TTTCAGAAGGCCTTATTTTTGGACATTTTATAGTGTAGTACTGTTGGCTCGATATTTGAATATTCAGCTACGACAGT
TTTGGACTGTTTAGTAGTCTTTGTTTTATGTTTTTAAATACAGAAATTGCTTCACAAATTTGTACCACATGGTAATT
CTAAGACTTGTTCTTTACCCATGGAATGTAATATTTTTGCAAAGATGGACTACTTCACAAATGGTTATAAAGTCATA
TCCACTTCTTCCACAATGACCACAGCAAATGACCCAAGCATGAACTAAAGAAGAG

FIGURE 81

CGCCAGC**ATG**CTGCCGGAGCAGCTCTACTTCCTGCAGAGCCCCCGGAGGAGGAGCCCGAATACCACCCCGACGCCT
CAGCCCAAGAATTAAATGTGAGGGAGTCTGACGTAAGAGTTTGTGATGAGTCATCATGTAAATATGGAGGAGTCTGT
AAAGAAGATGGAGATGGTTTGAAATGTGCATGCCAATTTCAGTGCCATACAAATTATATTCCTGTCTGTGGATCAAA
TGGGGACACTTATCAAAATGAATGCTTTCTCAGAAGGGCTGCTTGTAAGCACCAGAAAGAGATAACAGTAATAGCAA
GAGGACCATGCTACTCTGATAATGGATCTGGATCTGGAGAAGGAGAAGAGGAAGGGTCAGGGGCAGAAGTTCACAGA
AAACACTCCAAGTGTGGACCCTGCAAATATAAAGCTGAGTGTGATGAAGATGCAGAAAATGTTGGGTGTGTATGTAA
TATAGATTGCAGTGGATACAGTTTTAATCCTGTGTGTGCTTCTGATGGGAGTTCCTATAACAATCCCTGTTTTGTTC
GAGAAGCATCTTGGTATAAAGCAAGAACAAAATTGATATAAGGCATCTTGGTCATTGCACAGATACAGATGACACTAGT
TTGTTGGGGAAAGAAAGATGATGGACTACAATATCGACCAGATGTGAAAGATGCTAGTGATCAAAGAGAAGATGTTTA
TATTGGAAACCACATGCCTTGCCCTGAAAACCTCAATGGTTACTGCATCCATGGAAAATGTGAATTCATCTATTCTA
CTCAGAAGGCTTCTTGTAGATGTGAATCTGGCTACACTGGACAGCACTGTGAAAAGACAGACTTTAGTATTCTCTAT
GTAGTGCCAAGTAGGCAAAAGCTCACTCATGTTCTTATTGCAGCAATTATTGGAGCTGTACAGATTGCCATCATAGT
AGCAATTGTAATGTGCATAACAAGAAAATGCCCCAAAAACAATAGAGGACGTCGACAGAAGCAAAACCTAGGTCATT
TTACTTCAGATACGTCATCCAGAATGGTT**TAA**ACTGATGACTTTTATATGTACACTGACCATGTGATGTACATTTAT
TATGTCTTTTTTTAAAGAATGGAAATATTTATTTCAGAGGCCTTATTTTTGGACATTTTTAGTGTAGTACTGTTGGC
TCGTATTTAGAATATTCAGCTACGACAGTTTTGGACTGTTTAGTAGTCTTTGTTTTATGTTTTTAAATACAGAAATT
GCTTTCACAAATTTGTACCACATGGTAATTCTAAGACTTGTTCTTTACCCATGGAATGTAATATTTTTGCAAAGATG
GACTACTTCACAAATGGTTATAAAGTCATATCCACTTCTTCCACAATGACCACAGCAAATGACCAAGCATGAACTAA
AGGTAAAGATGTTTACAGATTACTTTTCTTACAAAAAAAATCTAGAAGACACTGTGTTTAAATAGATATTTAAATGT
TTTTGAGATTTAGTAACTGATTTTTTAGACACTGCCTATCGCATGAACTGTAAAGCTGTGTGTATTAGGTGTAAAAT
ATTTATAAGATATATGGACTGGGGAATTTGATTATTCCTCCCTTTGAAAAAATAGTCCTAATAATTTGAACAAATAT
GTTAGTAATGATGGAACAGATCAATGAAAAGTAGATATAGATATTGTGAAAATAGGCTGTTTAACAAACAGATTGGA
ATAAAGCCTATTCTACCAGTTAAACTACTTTAATACACATTCATTTTTAAAGAAAATGTTTGTTTTAACATAAATAA
ACAAATCGTATCAGTGTTTGTGAATAAAATACAAAAATGATTGTTAATGATTGGTGCTCTTAAAGTGAGCTTAAAAT
TTATCCAAGACGTATATCCAAATTTGTCCTGTAGTAATAGATTAATATTCATAGATTGTTGGTGTTTAAAGATCTGA
AGTGTGAGTAGAATGTATTCAGCTGTTTAACATGTAGTTTAGATATTCAAAAGTATGCATGTAGAATTTAAAGAATA
TGTTAAAAATTATTAATTTTAATATTTTGTTTGGAAAAGCATGTTTATAATATAATGTTTTCACTATAAAGAAAAAAA
AAAAAAAAAAA

FIGURE 82

MPTVDDILEQVGESGWFQKQAFLILCLLSAAFAPICVGIVFLGFTPDHHCQSPGVAELSQRCGWSPAEELNYTVPGL
GPAGEAFLGQCRRYEVDWNQSALSCVDPLASLATNRSHLPLGPCQDGWVYDTPGSSIVTEFNLVCADSWKLDLFQSC
LNAGFLFGSLGVGYFADRFGRKLCLLGTVLVNAVSGVLMAFSPNYMSMLLFRLLQGLVSKGNWMAGYTLITEFVGSG
SRRTVAIMYQMAFTVGLVALTGLAYALPHWRWLQLAVSLPTFLFLLYYWCVPESPRWLLSQKRNTEAIKIMDHIAQK
NGKLPPADLKMLSLEEDVTEKLSPSFADLFRTPRLRKRTFILMYLWFTDSVLYQGLILHMGATSGNLYLDFLYSALV
EIPGAFIALITIDRVGRIYPMAMSNLLAGAACLVMIFISPDLHWLNIIIMCVGRMGITIAIQMICLVNAELYPTFVR
NLGVMVCSSLCDIGGIITPFIVFRLREVWQALPLILFAVLGLLAAGVTLLLPETKGVALPETMKDAENLGRKAKPKE
NTIYLKVQTSEPSGT

Transmembrane domains.
amino acids 19-39, 175-195, 198-218, 235-255, 264-284, 350-370, 377-397,
409-429, 457-477, 492-512

N-glycosylation sites.
amino acids 71-74, 96-99, 112-115

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 293-296, 344-347

Tyrosine kinase phosphorylation site.
amino acids 536-544

N-myristoylation sites.
amino acids 76-81, 190-195, 215-220, 231-236, 253-258, 369-374, 389-394,
441-446, 465-470, 503-508

Amidation sites.
amino acids 173-176, 531-534

Sugar (and other) transporter homology.
amino acids 118-528

FIGURE 83

MRSDKSALVFLLLQLFCVGCGFCGKVLVWPCDMSHWLNVKVILEELIVRGHEVTVLTHSKPSLIDYRKPSALKFEVV
HMPQDRTEENEIFVDLALNVLPGLSTWQSVIKLNDFFVEIRGTLKMMCESFIYNQTLMKKLQETNYDVMLIDPVIPC
GDLMAELLAVPFVLTLRISVGGNMERSCGKLPAPLSYVPVPMTGLTDRMTFLERVKNSMLSVLFHFWIQDYDYHFWE
EFYSKALGRPTTLCETVGKAEIWLIRTYWDFEFPQPYQPNFEFVGGLHCKPAKALPKEMENFVQSSGEDGIVVFSLG
SLFQNVTEEKANIIASALAQIPQKVLWRYKGKKPSTLGANTRLYDWIPQNDLLGHPKTKAFITHGGMNGIYEAIYHG
VPMVGVPIFGDQLDNIAHMKAKGAAVEINFKTMTSEDLLRALRTVITDSSYKENAMRLSRIHHDQPVKPLDRAVFWI
EFVMRHKGAKHLRSAAHDLTWFQHYSIDVIGFLLTCVATAIFLFTKCFLFSCQKFNKTRKIEKRE

Signal sequence.
amino acids 1-21

Transmembrane domain.
amino acids 489-509

N-glycosylation sites.
amino acids 131-134, 313-316, 518-521

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 67-70, 340-343

Tyrosine kinase phosphorylation sites.
amino acids 122-131, 136-144

N-myristoylation sites.
amino acids 19-24, 276-281, 373-378, 377-382

Amidation site.
amino acids 338-341

FIGURE 84

MQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALVLAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGP
VGTADPAEKTPLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGS
LELLLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLYH
LGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRV
FYRPYLQIFGQIPQEDMDVALMEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMF
SYTFGKVQGNSDLYWKAQVTASSGNSTLGPRWPRPLSSSPTCASCSGNCAGDPGAPSRPPRPSSISGFTFLRKPSGS
ANVGIGA

Transmembrane domains.
amino acids 26-46, 118-138, 203-223, 233-253, 269-289, 358-378

N-glycosylation sites.
amino acids 334-337, 410-413

Glycosaminoglycan attachment site.
amino acids 176-179

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 457-460

N-myristoylation sites.
amino acids 99-104, 102-107, 106-111, 119-124, 127-132, 174-179, 177-182, 179-
184, 180-185, 347-352, 351-356, 432-437

Amidation site.
amino acids 106-109

Leucine zipper pattern.
amino acids 211-232

FIGURE 85

MTLWNGVLPFYPQPRHAAGFSVPLLIVILVFLALAASFLLILPGIRGHSRWFWLVRVLLSLFIGAEIVAVHFSAEWF
VGTVNTNTSYKAFSAARVTARVGLLVGLEGINITLTGTPVHQLNETIDYNEQFTWRLKENYAAEYANALEKGLPDPV
LYLAEKFTPSSPCGLYHQYHLAGHYASATLWVAFCFWLLSNVLLSTPAPLYGGLALLTTGAFALFGVFALASISSVP
LCPLRLGSSALTTQYGAAFWVTLATGVLCLFLGGAVVSLQYVRPSALRTLLDQSAKDCSQERGGSPLILGDPLHKQA
ALPDLKCITTNL

Transmembrane domains.
amino acids 19-39, 54-74, 94-114, 178-198, 211-231, 251-271

N-glycosylation sites.
amino acids 84-87, 109-112, 121-124

Tyrosine kinase phosphorylation site.
amino acids 148-157

N-myristoylation sites.
amino acids 79-84, 100-105, 107-112, 265-270

Leucine zipper pattern.
amino acids 197-218

FIGURE 86

MERLWGLFQRAQQLSPRSSQTVYQRVEGPRKGHLEEEEEDGEEGAETLAHFCPMELRGPEPLGSRPRQPNLIPWAAA
GRRAAPYLVLTALLIFTGAFLLGYVAFRGSCQACGDSVLVVSEDVNYEPDLDFHQGRLYWSDLQAMFLQFLGEGRLE
DTIRQTSLRERVAGSAGMAALTQDIRAALSRQKLDHVWTDTHYVGLQFPDPAHPNTLHWVDEAGKVGEQLPLEDPDV
YCPYSAIGNVTGELVYAHYGRPEDLQDLRARGVDPVGRLLLVRVGVISFAQKVTNAQDFGAQGVLIYPEPADFSQDP
PKPSLSSQQAVYGHVHLGTGDPYTPGFPSFNQTQFPPVASSGLPSIPAQPISADIASRLLRKLKGPVAPQEWQGSLL
GSPYHLGPGPRLRLVVNNHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDAWGPGAAKSAVGTAILLELVRTFSSMVS
NGFRPRRSLLFISWDGGDFGSVGSTEWLEGYLSVLHLKAVVYVSLDNAVLGDDKFHAKTSPLLTSLIESVLKQVDSP
NHSGQTLYEQVVFTNPSWDAEVIRPLPMDSSAYSFTAFVGVPAVEFSFMEDDQAYPFLHTKEDTYENLHKVLQGRLP
AVAQAVAQLAGQLLIRLSHDRLLPLDFGRYGDVVLRHIGNLNEFSGDLKARGLTLQWVYSARGDYIRAAEKLRQEIY
SSEERDERLTRMYNVRIMRVEFYFLSQYVSPADSPFRHIFMGRGDHTLGALLDHLRLLRSNSSGTPGATSSTGFQES
RFRRQLALLTWTLQGAANALSGDVWNIDNNF

Transmembrane domain.
amino acids 79-99

N-glycosylation sites.
amino acids 240-243, 339-342, 540-543, 754-757

N-myristoylation sites.
amino acids 106-111, 239-244, 382-387, 415-420, 439-444, 478-483, 482-487, 757-762, 760-765, 785-790

Amidation site.
amino acids 77-80

Cell attachment sequences.
amino acids 678-680, 736-738

PA domain.
amino acids 236-335

FIGURE 87

MVELMFPLLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKELAQRGARVYLACRDVEK
GELVAKEIQTTTGNQQVLVRKLDLSDTKSIRAFAKGFLAEEKHLHVLINNAGVMMCPYSKTADGFEMHIGVNHLGHF
LLTHLLLEKLKESAPSRIVNVSSLAHHLGRIHFHNLQGEKFYNAGLAYCHSKLANILFTQELARRLKGSGVTTYSVH
PGTVQSELVRHSSFMRWMWWLFSFFIKTPQQGAQTSLHCALTEGLEILSGNHFSDCHVAWVSAQARNETIARRLWDV
SCDLLGLPID

Signal sequence.
amino acids 1-20

Transmembrane domain.
amino acids 32-52

N-glycosylation sites.
amino acids 174-177, 298-301

N-myristoylation sites.
amino acids 31-36, 48-53, 199-204, 222-227, 233-238, 263-268, 281-286

Short chain dehydrogenase.
amino acids 40-303

FIGURE 88

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNLVTFGHSLSKTCSPACPI
PEGVNVGVASMGISCCQSFLCNFSAADGGLRASVTLLGAGLLLSLLPALLRFGP

Signal sequence.
amino acids 1-20

Transmembrane domain.
amino acids 108-128

N-glycosylation site.
amino acids 99-102

N-myristoylation sites.
amino acids 14-19, 58-63, 80-85, 89-94, 105-110, 106-111, 117-122

u-PAR/Ly-6 domain.
amino acids 21-100

FIGURE 89

MLLLKKHTEDISSVYEIRERLGSGAFSEVVLAQERGSAHLVALKCIPKKALRGKEALVENEIAVLRRISHPNIVALE
DVHESPSHLYLAMELVTGGELFDRIMERGSYTEKDASHLVGQVLGAVSYLHSLGIVHRDLKPENLLYATPFEDSKIM
VSDFGLSKIQAGNMLGTACGTPGYVAPELLEQKPYGKAVDVWALGVISYILLCGYPPFYDESDPELFSQILRASYEF
DXPFWDDISESGKDFIRHLLERDLQKRFTCQQALRDLWIFWDTGFGRDILGFVSEQIRKNFAWTHWKRAFNATLFLR
HIRKLGQIPEGEGASEQGMXRHSHXGLRAGQPPKW

N-glycosylation site.
amino acids 302-305

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 5-8, 66-69, 257-260

Tyrosine kinase phosphorylation site.
amino acids 101-109

N-myristoylation sites.
amino acids 118-123, 166-171, 170-175, 334-339

Serine/Threonine protein kinases active-site signature.
amino acids 132-145

Protein kinase domain.
amino acids 15-270

FIGURE 90

MEELQDDYEDMMEENLEQEEYEDPDIPESQMEEPAAHDTEATATDYHTTSHPGTHEVYVELQELVMDEKNQELRWME
AARWVQLEENLGENGAWGRPHLSHLTFWSLLELRRVFTKGTVLLDLQETSLAGVANQLLDRFIFEDQIRPQDREELL
RALLLKHSHAGELEALGGVKPAVLTRSGDPSQPLLPQHSSLETQLFCEQGDGGTEGHSPSGILEKIPPDSEATLVLV
GRADFLEQPVLGFVRLQEAAELEAVELPVPIRFLFVLLGPEAPHIDYTQLGRAAATLMSERVFRIDAYMAQSRGELL
HSLEGFLDCSLVLPPTDAPSEQALLSLVPVQRELLRRRYQSSPAKPDSSFYKGLDLNGGPDDPLQQTGQLFGGLVRD
IRRRYPYYLSDITDAFSPQVLAAVIFIYFAALSPAITFGGLLGEKTRNQMGVSELLISTAVQGILFALLGAQPLLVV
GFSGPLLVFEEAFFSFCETNGLEYIVGRVWIGFWLILLVVLVVAFEGSFLVRFISRYTQEIFSFLISLIFIYETFSK
LIKIFQDHPLQKTYNYNVLMVPKPQGPLPNTALLSLVLMAGTFFFAMMLRKFKNSSYFPGKLRRVIGDFGVPISILI
MVLVDFFIQDTYTQKLSVPDGFKVSNSSARGWVIHPLGLRSEFPIWMMFASALPALLVFILIFLESQITTLIVSKPE
RKMVKGSGFHLDLLLVVGMGGVAALFGMPWLSATTVRSVTHANALTVMGKASTPGAAAQIQEVKEQRISGLLVAVLV
GLSILMEPILSRIPLAVLFGIFLYMGVTSLSGIQLFDRILLLFKPPKYHPDVPYVKRVKTWRMHLFTGIQIICLAVL
WVVKSTPASLALPFVLILTVPLRRVLLPLIFRNVELQCLDADDAKATFDEEEGRDEYDEVAMPV

Transmembrane domains.
amino acids 401-421, 445-465, 487-507, 519-539, 564-584, 602-622, 663-683, 700-
720, 761-781, 793-813, 831-851, 854-874

N-glycosylation site.
amino acids 593-596, 642-645

Tyrosine kinase phosphorylation site.
amino acids 817-824

N-myristoylation sites.
amino acids 172-177, 206-211, 361-366, 376-381, 424-429, 448-453, 609-614, 713-
718, 763-768

Anion exchangers family signature 1.
amino acids 379-390

Anion exchangers family signature 2.
amino acids 526-540

FIGURE 91

```
MVAWVVSTGVAIACCAAVYYLAEYNLEFLKTHSNPGAVLLLPFVVSCINLAVPCIYSMFRLVERYEMPRHEVYVLLI
RNIFLKISIIGILCYYWLNTVALSGEECWETLIGQDIYRLLLMDFVFSLVNSFLGEFLRRIIGMQLITSLGLQEFDI
ARNVLELIYAQTLVWIGIFFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPSKAWRASQMMTFFIFLLFFPSFTGVLC
TLAITIWRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRPGYLWVVWIYRNLIGSVHFFFILTLIVLIITYLYWQITE
GRKIMIRLLHEQIINEGKDKMFLIEKLIKLQDMEKKANPSSLVLERREVEQQGFLHLGEHDGSLDLRSRRSVQEGNPRA
```

Transmembrane domains.
amino acids 1-16, 34-54, 75-95, 114-134, 159-179, 175-195, 213-233, 250-270,
281-301

N-glycosylation site.
amino acids 194-197

Tyrosine kinase phosphorylation site.
amino acids 156-163

N-myristoylation sites.
amino acids 9-14, 228-233

Amidation site.
amino acids 308-311

FIGURE 92

MERRRLWGSIQSRYISMSVWTSPRRLVELAGQSLLKDEALAIAALELLPRELFPPLFMAAFDGRHSQTLKAMVQAWP
FTCLPLGVLMKGQHLHLETFKAVLDGLDVLLAQEVRPRRWKLQVLDLRKNSHQDFWTVWSGNRASLYSFPEPEAAQP
MTKKRKVDGLSTEAEQPFIPVEVLVDLFLKEGACDELFSYLIEKVKRKKNVLRLCCKKLKIFAMPMQDIKMILKMVQ
LDSIEDLEVTCTWKLPTLAKFSPYLGQMINLRRLLLSHIHASSYISPEKEEQYIAQFTSQFLSLQCLQALYVDSLFF
LRGRLDQLLRHVMNPLETLSITNCRLSEGDVMHLSQSPSVSQLSVLSLSGVMLTDVSPEPLQALLERASATLQDLVF
DECGITDDQLLALLPSLSHCSQLTTLSFYGNSISISALQSLLQHLIGLSNLTHVLYPVPLESYEDIHGTLHLERLAY
LHARLRELLCELGRPSMVWLSANPCPHCGDRTFYDPEPILCPCFMPN

N-glycosylation site.
amino acids 435-438

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 125-128

N-myristoylation sites.
amino acids 8-13, 138-143, 257-262, 358-363, 415-418

FIGURE 93

MTQRSIAGPICNLKFVTLLVALSSELPFLGAGVQLQDNGYNGLLIAINPQVPENQNLISNIKEMITEASFYLFNATK
RRVFFRNIKILIPATWKANNNSKIKQESYEKANVIVTDWYGAHGDDPYTLQYRGCGKEGKYIHFTPNFLLNDNLTAG
YGSRGRVFVHEWAHLRWGVFDEYINDKPFYINGQNQIKVTRCSSDITGIFVCEKGPCPQENCIISKLFKEGCTFIYN
STQNATASIMFMQSLSSVVEFCNASTHNQEAPNLQNQMCSLRSAWDVITDSADFHHSFPMNGTELPPPPTFSLVQAG
DKVVCLVLDVSSKMAEADRLLQLQQAAEFYLMQIVEIHTFVGIASFDSKGEIRAQLHQINSNDDRKLLVSYLPTTVS
AKTDISICSGLKKGFEVVEKLNGKAYGSVMILVTSGDDKLLGNCLPTVLSSGSTIHSIALGSSAAPNLEELSRLTGG
LKFFVPDISNSNSMIDAFSRISSGTGDIFQQHIQLESTGENVKPHHQLKNTVTVDNTVGNDTMFLVTWQASGPPEII
LFDPDGRKYYTNNFITNLTFRTASLWIPGTAKPGHWTYTLNNTHHSLQALKVTVTSRASNSAVPPATVEAFVERDSL
HFPHPVMIYANVKQGFYPILNATVTATVEPETGDPVTLRLLDDGAGADVIKNDGIYSRYFFSFAANGRYSLKVHVNH
SPSISTPAHSIPGSHAMYVPGYTANGNIQMNAPRKSVGRNEEERKWGFSRVSSGGSFSVLGVPAGPHPDVFPPCKII
DLEAVKVEEELTLSWTAPGEDFDQGQATSYEIRMSKSLQNIQDDFNNAILVNTSKRNPQQAGIREIFTFSPQISTNG
PEHQPNGETHESHRIYVAIRAMDRNSLQSAVSNIAQAPLFIPPNSDPVPARDYLILKGVLTAMGLIGIICLIIVVTH
HTLSRKKRADKKENGTKLL

Signal sequence.
amino acids 1-31

Transmembrane domain.
amino acids 903-923

N-glycosylation sites.
amino acids 74-77, 97-100, 150-153, 231-234, 235-238, 254-257, 292-295, 522-525,
556-559, 580-583, 637-640, 822-825, 938-941

Glycosaminoglycan attachment site.
amino acids 485-488

N-myristoylation sites.
amino acids 42-47, 202-207, 395-400, 747-752, 795-800, 905-910

Amidation site.
amino acids 544-547

FIGURE 94

MALNSGSPPAIGPYYENHGYQPENPYPAQPTVVPTVYEVHPAQYYPSPVPQYAPRVLTQASNPVVCTQPKSPSGTVC
TSKTKKALCITLTLGTFLVGAALAAGLLWKFMGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGP
NFILQMYSSQRKSWHPVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDAC
SSKAVVSLRCLACGVNLNSSRQSRIVGGESALPGAWPWQVSLHVQNVHVCGGSIITPEWIVTAAHCVEKPLNNPWHW
TAFAGILRQSFMFYGAGYQVQKVISHPNYDSKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWG
ATEEKGKTSEVLNAAKVLLIETQRCNSRYVYDNLITPAMICAGFLQGNVDSCQGDSGGPLVTSNNNIWWLIGDTSWG
SGCAKAYRPGVYGNVMVFTDWIYRQMKANG

Transmembrane domain.
amino acids 86-106

N-glycosylation sites.
amino acids 213-216, 249-252

Glycosaminoglycan attachment site.
amino acids 382-385

N-myristoylation sites.
amino acids 6-11, 74-79, 97-102, 110-115, 245-250, 258-263, 432-437, 462-467, 464-469, 472-477

ATP/GTP-binding site motif A (P-loop).
amino acids 386-393

Serine proteases, trypsin family, histidine active site.
amino acids 292-297

Trypsin homology.
amino acids 256-484

FIGURE 95

MARSLLLPLQILLLSLALETAGEEAQGDKIIDGAPCARGSHPWQVALLSGNQLHCGGVLVNERWVLTAAHCKMNEYT
VHLGSDTLGDRRAQRIKASKSFRHPGYSTQTHVNDLMLVKLNSQARLSSMVKKVRLPSRCEPPGTTCTVSGWGTTTS
PDVTFPSDLMCVDVKLISPQDCTKVYKDLLENSMLCAGIPDSKKNACNGDSGGPLVCRGTLQGLVSWGTFPCGQPND
PGVYTQVCKFTKWINDTMKKHR

Signal sequence.
amino acids 1-15

N-glycosylation site.
amino acids 246-249

Glycosaminoglycan attachment site.
amino acids 147-150

N-myristoylation sites.
amino acids 33-38, 57-62, 141-146, 192-197, 213-218, 222-227

Serine proteases, trypsin family, histidine active site.
amino acids 66-71

Trypsin homology.
amino acids 30-245

FIGURE 96

MESRKDITNQEELWKMKPRRNLEEDDYLHKDTGETSMLKRPVLLHLHQTAHADEFDCPSELQHTQELFPQWHLPIKI
AAIIASLTFLYTLLREVIHPLATSHQQYFYKIPILVINKVLPMVSITLLALVYLPGVIAAIVQLHNGTKYKKFPHWL
DKWMLTRKQFGLLSFFFAVLHAIYSLSYPMRRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEIYVSLGIVGLAILA
LLAVTSIPSVSDSLTWREFHYIQSKLGIVSLLLGTIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLIFKSIL
FLPCLRKKILKIRHGWEDVTKINKTEICSQL

Transmembrane domains.
amino acids 72-92, 116-136, 158-178, 218-238, 254-274, 291-311

N-glycosylation sites.
amino acids 143-146, 331-334

Tyrosine kinase phosphorylation site.
amino acids 19-27

N-myristoylation sites.
amino acids 133-138, 265-270

FIGURE 97

MDSSPSLPLIRTPESSLHEALDQCMTALDLFLTNQFSEALSYLKPRTKESMYHSLTYATILEMQAMMTFDPQDILLA
GNMMKEAQMLCQRHRRKSSVTDSFSSLVNRPTLGQFTEEEIHAEVCYAECLLQRAALTFLQGSSHGGAVRPRALHDP
SHACSCPPGPGRQHLFLLQDENMVSFIKGGIKVRNSYQTYKELDSLVQSSQYCKGENHPHFEGGVKLGVGAFNLTLS
MLPTRILRLLEFVGFSGNKDYGLLQLEEGASGHSFRSVLCVMLLLCYHTFLTFVLGTGNVNIEEAEKLLKPYLNRYP
KGAIFLFFAGRIEVIKGNIDAVSDGGPGRGWGSLGVSQTSRKSGTCDILRDRIDWGRGGGQERTNQRAGAGEALLAE
QPGKTREEEAFVVPGILTGRYRTAALQWREVEGGA

Transmembrane domain.
amino acids 267-287

N-glycosylation site.
amino acids 227-230

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 92-95, 93-96

Tyrosine kinase phosphorylation site.
amino acids 46-52

N-myristoylation sites.
amino acids 139-144, 218-223, 325-330, 343-348, 368-373, 377-382, 400-405

Eukaryotic putative RNA-binding region RNP-1 signature.
amino acids 309-316

FIGURE 98

MADAEVIILPKKHKKKKERKSLPEEDVAEIQHAEEFLIKPESKVAKLDTSQWPLLLKNFDKLNVRTTHYTPLACGSN
PLKREIGDYIRTGFINLDKPSNPSSHEVVAWIRRILRVEKTGHSGTLDPKVTGCLIVCIERATRLVKSQQSAGKEYV
GIVRLHNAIEGGTQLSRALETLTGALFQRPPLIAAVKRQLRVRTIYESKMIEYDPERRLGIFWVSCEAGTYIRTLCV
HLGLLLGVGGQMQELRRVRSGVMSEKDHMVTMHDVLDAQWLYDNHKDESYLRRVVYPLEKLLTSHKRLVMKDSAVNA
ICYGAKIMLPGVLRYEDGIEVNQEIVVITTKGEAICMAIALMTTAVISTCDHGIVAKIKRVIMERDTYPRKWGLGPK
ASQKKLMIKQGLLDKHGKPTDSTPATWKQEYVDYSESAKKEVVAEVVKAPQVVAEAAKTAKRKRESESESDETPPAA
PQLIKKEKKKSKKDKKAKAGLESGAEPGDGDSDTTKKKKKKKKAKEVELVSE

Transmembrane domain.
amino acids 220-240

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 448-451, 470-473

Tyrosine kinase phosphorylation site.
amino acids 368-376

N-myristoylation sites.
amino acids 166-171, 234-239, 326-331, 482-487, 486-491

TruB family pseudouridylate synthase homology.
amino acids 107-247

PUA domain.
amino acids 295-370

FIGURE 99

MALRICVTYTPALPIGLCTRCCLCLEQSPSWCHCLRGVSFLTFHLHQSVPLGDRDSLLMFTRQAGHFVEGSKAGRSR
GRLCLSQALRVAVRGAFVSLWFAAGAGDRERNKGDKGAQTGAGLSQEAEDVDVSRARRVTDAPQGTLCGTGNRNSGS
QSARAVGVAHLGEAFRVGVEQAISSCPEEVHGRHGLSMEIMWARMDVALRSPGRGLLAGAGALCVTLAESSCPDYER
GRRACLTLHRHPTPHCSTWGLPLRVAGSWLTVVTVEALGGWRMGVRRTGQVGPTMHPPPVSGASPLLLHHLLLLLLI
IILTC

Signal sequence.
amino acids 1-34

Transmembrane domains.
amino acids 80-100, 250-270, 293-312

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 134-137

N-myristoylation sites.
amino acids 70-75, 92-97, 114-119, 118-123, 142-147, 148-153, 153-158, 172-177,
209-214, 258-263, 275-280

Amidation site.
amino acids 231-234

Eukaryotic putative RNA-binding region RNP-1 signature.
amino acids 36-43

FIGURE 100

MGIVANFQELAVPVVHDGGALLAFVCGVVYTLLQSIISYKSCPQWNSLSTCHIRMVISAVSCAAVIPMIVCASLISI
TKLEWNPREKDYVYHVVSAICEWTVAFGFIFYFLTFIQDFQSVTLRISTEINGDI

Transmembrane domains.
amino acids 10-30, 53-73, 93-113

Tyrosine kinase phosphorylation site.
amino acids 85-91

N-myristoylation sites.
amino acids 2-7, 19-24, 27-32

FIGURE 101

MPRQLSAAAALFASLAVILHDGSQMRAKAFPETRDYSQPTAAATVQDIKKPVQQPAKQAPHQTLAARFMDGHITFQT
AATVKIPTTTPATTKNTATTSPITYTLVTTQATPNNSHTAPPVTEVTVGPSLAPYSLPPTITPPAHTAGTSSSTVSH
TTGNTTQPSNQTTLPATLSIALHKSTTGQKPDQPTHAPGTTAAAHNTTRTAAPASTVPGPTLAPQPSSVKTGIYQVL
NGSRLCIKAEMGIQLIVQDKESVFSPRRYFNIDPNATQASGNCGTRKSNLLLNFQGGFVNLTFTKDEESYYISEVGA
YLTVSDPETVYQGIKHAVVMFQTAVGHSFKCVSEQSLQLSAHLQVKTTDVQLQAFDFEDDHFGNVDECSSDYTIVLP
VIGAIVVGLCLMGMGVYKIRLRCQSSGYQRI

Signal sequence.
amino acids 1-16

Transmembrane domain.
amino acids 380-400

N-glycosylation sites.
amino acids 112-115, 158-161, 164-167, 200-203, 232-235, 266-269, 291-294

Tyrosine kinase phosphorylation sites.
amino acids 28-36, 296-302

N-myristoylation sites.
amino acids 146-151, 193-198, 233-238, 272-277, 275-280, 287-292, 307-312, 321-326

Lysosome-associated membrane glycoprotein homology.
amino acids 66-416

FIGURE 102

MTEEPIKEILGAPKAHMAATMEKSPKSEVVITTVPLVSEIQLMAATGGTELSCYRCIIPFAVVVFIAGIVVTAVAYS
FNSHGSIISIFGLVVLSSGLFLLASSALCWKVRQRSKKAKRRESQTALVANQRSLFA

Transmembrane domains.
amino acids 54-74, 84-104

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 118-121

N-myristoylation sites.
amino acids 11-16, 48-53, 68-73, 82-87

FIGURE 103

MLSLLHASTLAVLGALCVYGAGHLEQPQISSTKTLSKTARLECVVSGITISATSVYWYRERPGEVIQFLVSISYDGT
VRKESGIPSGKFEVDRIPETSTSTLTIHNVEKQDIATYYCALWEVRLANQELGKKIKVFGPGTKLIITDKQLDADVS
PKPTIFLPSIAETKLQKAGTYLCLLEKFFPDVIKIHWQEKKSNTILGSQEGNTMKTNDTYMKFSWLTVPEKSLDKEH
RCIVRHENNKNGVDQEIIFPPIKTDVITMDPKDNCSKDANDTLLLQLTNTSAYYTYLLLLLKSVVYFAIITCCLLRR
TAFCCNGEKS

Signal sequence.
amino acids 1-17

Transmembrane domain.
amino acids 282-302

N-glycosylation sites.
amino acids 211-214, 265-268, 271-274, 280-283

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 79-82

N-myristoylation sites.
amino acids 47-52, 83-88, 173-178, 201-206

Amidation site.
amino acids 129-132

Immunoglobulin domain.
amino acids 36-119

FIGURE 104

MNYSLHLAFVCLSLFTERMCIQGSQFNVEVGRSDKLSLPGFENLTAGYNKFLRPNFGGEPVQIALTLDIASISSISE
SNMDYTATIYLRQRWMDQRLVFEGNKSFTLDARLVEFLWVPDTYIVESKKSFLHEVTVGNRLIRLFSNGTVLYALRI
TTTVACNMDLSKYPMDTQTCKLQLESWGYDGNDVEFTWLRGNDSVRGLEHLRLAQYTIERYFTLVTRSQQETGNYTR
LVLQFELRRNVLYFILETYVPSTFLVVLSWVSFWISLDSVPARTCIGVTTVLSMTTLMIGSRTSLPNTNCFIKAIDV
YLGICFSFVFGALLEYAVAHYSSLQQMAAKDRGTTKEVEEVSITNIINSSISSFKRKISFASIEISSDNVDYSDLTM
KTSDKFKFVFREKMGRIVDYFTIQNPSNVDHYSKLLFPLIFMLANVFYWAYYMYF

Signal sequence.
amino acids 1-16

Transmembrane domains.
amino acids 306-326, 415-435

N-glycosylation sites.
amino acids 2-5, 43-46, 102-105, 145-148, 196-199, 228-231, 356-259

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 364-367

Tyrosine kinase phosphorylation sites.
amino acids 175-183, 221-229

N-myristoylation sites.
amino acids 23-28, 291-296, 311-316

Neurotransmitter-gated ion-channels signature.
amino acids 160-174

Neurotransmitter-gated ion-channel ligand binding domain.
amino acids 37-242

Neurotransmitter-gated ion-channel transmembrane region.
amino acids 249-434

FIGURE 105

MILVLLCAFLIPCPPRDLHSTWSRHLGSQGGGDLSPLELADVNGDGLRDVLLSFVMSRNGSAVGVSRPAANLVCLSG
MNGSTLWSSLLPEEARDITCLELMPGSLAETICLVTGTHKMLSAFNATSGKAIWTLNPNYLSNGTLAAPVVVLPDLD
EDGVRDLVVLAIGELQPDLCFLLVSGRTGNPVGRPVKYNIVGVGNLIGPQVYITTNGAVYILFGFGNIQAVALRDIF
VQAQNRDSSPPSLQIEEPEWEKRRSINLSELIDVYSDGVELLQMVKAPDSNCSNLLITTRQSLVLLRGQNLTPYWAL
RLQGLRSQPTPGYFTDDQTLDFLLQIQDGVGMKKMMVVDGDSGSIVWSYRAPCHMKETPATSAVTSDQKSVFLFWAE
GLSAASPNSDIILGTEPPSLHHLYLLHPAFPSILLDLANTTGTVTASEVGINDLWKDAFYVTRTTGPSSEGHPAALV
VSKLSLRWALMEGQMAQLQESTPKIGRGELRRFLSRIKFVEAPYEI

Signal sequence.
amino acids 1-15

Transmembrane domain.
amino acids 202-222

N-glycosylation sites.
amino acids 59-62, 79-82, 123-126, 140-143, 258-261, 282-285, 424-427

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 253-256

Tyrosine kinase phosphorylation site.
amino acids 500-506

N-myristoylation sites.
amino acids 27-32, 60-65, 77-82, 183-188, 220-225, 386-391, 427-432

FIGURE 106

MSNKFLGTWKLVSSENFDDYMKALGVGLATRKLGNLAKPTVIISKKGDIITIRTESTFKNTEISFKLGQEFEETTAD
NRKTKSIVTLQRGSLNQVQRWDGKETTIKRKLVNGKMVAECKMKGVVCTRIYEKV

N-myristoylation sites.
amino acids 25-30, 122-127

Lipocalin/cytosolic fatty-acid binding protein homology.
amino acids 4-132

FIGURE 107

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLALALEPALASPAGAANF
LAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSW
TGFVGEDLVTIPKGFNTSFLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQ
MCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTL
LRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIYLRDENSSRSFRITILPQLYIQPMM
GAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQS
LSEPILWIVSYALMSVCGAILLVLIVLLLLPFRCQRRPRDPEVVNDESSLVRHRWK

Signal sequence.
amino acids 1-20

Transmembrane domain.
amino acids 471-491

N-glycosylation sites.
amino acids 170-173, 366-369

Glycosaminoglycan attachment site.
amino acids 88-91

N-myristoylation sites.
amino acids 59-64, 120-125, 152-157, 190-195, 198-203, 236-241, 241-246, 243-
248, 253-258, 285-290, 339-344, 386-391, 397-396, 438-443

Leucine zipper patterns.
amino acids 10-31, 197-218

Eukaryotic aspartyl protease homology.
amino acids 41-431

FIGURE 108

MAVKLGTLLLALALGLAQPASARRKLLVFLLDGFRSDYISDEALESLPGFKEIVSRGVKVDYLTPDFPSLSYPNYYT
LMTGRHCEVHQMIGNYMWDPTTNKSFDIGVNKDSLMPLWWNGSEPLWVTLTKAKRKVYMYYWPGCEVEILGVRPTYC
LEYKNVPTDINFANAVSDALDSFKSGRADLAAIYHERIDVEGHHYGPASPQRKDALKAVDTVLKYMTKWIQERGLQD
RLNVIIFSDHGMTDIFWMDKVIELNKYISLNDLQQVKDRGPVVSLWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRF
YYKKGKFVSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGIFLAFGPDFKSNFRAAPIRS
VDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALILLFLLA

Signal sequence.
amino acids 1-22

Transmembrane domain.
amino acids 423-439

N-glycosylation sites.
amino acids 100-103, 118-121, 341-344, 404-407

Tyrosine kinase phosphorylation sites.
amino acids 191-199, 251-258

N-myristoylation sites.
amino acids 148-153, 365-370

Amidation site.
amino acids 343-346

Phosphodiest Type I phosphodiesterase.
amino acids 5-381

FIGURE 109

MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTATTAASTTARKDIPVLP
KWVGDLPNGRVCP

Signal sequence.
amino acids 1-19

FIGURE 110

MFQTGGLIVFYGLLAQTMAQFGGLPVPLDQTLPLNVNPALPLSPTGLAGSLTNALSNGLLSGGLLGILENLPLLDIL
KPGGGTSGGLLGGLLGKVTSVIPGLNNIIDIKVTDPQLLELGLVQSPDGHRLYVTIPLGIKLQVNTPLVGASLLRLA
VKLDITAEILAVRDKQERIHLVLGDCTHSPGSLQISLLDGLGPLPIQGLLDSLTGILNKVLPELVQGNVCPLVNEVL
RGLDITLVHDIVNMLIHGLQFVIKV   .

Signal sequence.
amino acids 1-19

Transmembrane domain.
amino acids 134-154

N-myristoylation sites.
amino acids 46-51, 49-54, 58-63, 62-67, 66-71, 80-85, 81-86, 82-87, 85-90, 86-
91, 89-94, 202-207, 233-238

FIGURE 111

MEPPYSLTAHYDEFQEVKYVSRCGAGGARGASLPPGFPLGAARSVTGARSGLPRWNRREVCLLSGLVFAAGLCAILA
AMLALKYLGPVAAGGGACPEGCPERKAFARAARFLAANLDASIDPCQDFYSFACGGWLRRHAIPDDKLTYGTIAAIG
EQNEERLRRLLARPGGGPGGAAQRKVRAFFRSCLDMREIERLGPRPMLEVIEDCGGWDLGGAEERPGVAARWDLNRL
LYKAQGVYSAAALFSLTVSLDDRNSSRYVIRIDQDGLTLPERTLYLAQDEDSEKILAAYRVFMERVLSLLGADAVEQ
KAQEILQVEQQLANITVSEYDDLRRDVSSMYNKVTLGQLQKITPHLRWKWLLDQIFQEDFSEEEEVVLLATDYMQQV
SQLIRSTPHRVLHNYLVWRVVVVLSEHLSPPFREALHELAQEMEGSDKPQELARVCLGQANRHFGMALGALFVHEHF
SAASKAKVQQLVEDIKYILGQRLEELDWMDAETRAAARAKLQYMMVMVGYPDFLLKPDAVDKEYEFEVHEKTYFKNI
LNSIPFSIQLSVKKIRQEVDKSTWLLPPQALNAYYLPNKNQMVFPAGILQPTLYDPDFPQSLNYGGIGTIIGHELTH
GYDDWGGQYDRSGNLLHWWTEASYSRFLRKAECIVRLYDNFTVYNQRVNGKHTLGENIADMGVLKLAYHAYQKWVRE
HGPEHPLPRLKYTHDQLFFIAFAQNWCIKRRSQSIYLQVLTDKHAPEHYRVLGSVSQFEEFGRAFHCPKDSPMNPAH
KCSVW

Transmembrane domain.
amino acids 64-84

N-glycosylation sites.
amino acids 255-258, 322-325, 656-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 722-725

N-myristoylation sites.
amino acids 24-29, 26-31, 27-32, 40-45, 47-52, 65-70, 148-153, 169-174, 170-175,
237-242, 450-455, 604-609, 607-612

Prenyl group binding site (CAAX box).
amino acids 772-777

Neutral zinc metallopeptidases, zinc-binding region signature.
amino acids 609-618

Peptidase family M13.
amino acids 571-774

FIGURE 112

MGPSCPVFLSFTKLSLWWLLLTPAGGEEAKRPPPRAPGDPLSSPSPTALPQGGSHTETEDRLFKHLFRGYNRWARPV
PNTSDVVIVRFGLSIAQLIDVDEKNQMMTTNVWLKQEWSDYKLRWNPADFGNITSLRVPSEMIWIPDIVLYNNADGE
FAVTHMTKAHLFSTGTVHWVPPAIYKSSCSIDVTFFPFDQQNCKMKFGSWTYDKAKIDLEQMEQTVDLKDYWESGEW
AIVNATGTYNSKKYDCCAEIYPDVTYAFVIRRLPLFYTINLIIPCLLISCLTVLVFYLPSDCGEKITLCISVLLSLT
VFLLLITEIIPSTSLVIPLIGEYLLFTMIFVTLSIVITVFVLNVHHRSPSTHTMPHWVRGALLGCVPRWLLMNRPPP
PVELCHPLRLKLSPSYHWLESNVDAEEREVVVEEEDRWACAGHVAPSVGTLCSHGHLHSGASGPKAEALLQEGELLL
SPHMQKALEGVHYIADHLRSEDADSSVKEDWKYVAMVIDRIFLWLFIIVCFLGTIGLFLPPFLAGMI

Transmembrane domains.
amino acids 2-22, 264-284, 299-319, 330-350, 501-521

N-glycosylation sites.
amino acids 79-82, 129-132, 235-238

Tyrosine kinase phosphorylation site.
amino acids 468-475

N-myristoylation sites.
amino acids 25-30, 52-57, 89-94, 128-133, 238-243, 368-373, 434-439

Neurotransmitter-gated ion-channels signature.
amino acids 183-197

Neurotransmitter-gated ion-channel ligand binding domain.
amino acids 59-265

Neurotransmitter-gated ion-channel transmembrane domain.
amino acids 272-520

FIGURE 113

MLPPGTATLLTLLLAAGSLGQKPQRPRRPASPISTIQPKANFDAQQFAGTWLLVAVGSACRFLQEQGHRAEATTLHV
APQGTAMAVSTFRKLDGICWQVRQLYGDTGVLGRFLLQARGARGAVNVVVAETDYQSFAVLYLERAGQLSVKLYARS
LPVSDSVLSGFEQRVQEAHLTEDQIFYFPKYGFCEAADQFHVLDEVRR

Signal sequence.
amino acids 1-20

N-myristoylation sites.
amino acids 81-86, 118-123

Lipocalin/cytosolic fatty-acid binding protein.
amino acids 46-188

FIGURE 114

MRLTVLCAVCLLPGSLALPLPQEAGGMSELQWEQAQDYLKRFYLYDSETKNANSLEAKLKEMQKFFGLPITGMLNSR
VIEIMQKPRCGVPDVAEYSLFPNSPKWTSKVVTYRIVSYTRDLPHITVDRLVSKALNMWGKEIPLHFRKVVWGTADI
MIGFARGAHGDSYPFDGPGNTLAHAFAPGTGLGGDAHFDEDERWTDGSSLGINFLYAATHELGHSLGMGHSSDPNAV
MYPTYGNGDPQNFKLSQDDIKGIQKLYGKRSNSRKK

Signal sequence.
amino acids 1-17

N-myristoylation sites.
amino acids 67-72, 72-77, 173-178, 183-188, 201-206, 221-226

Amidation site.
amino acids 258-261

Neutral zinc metallopeptidases, zinc-binding region signature.
amino acids 211-220

Matrixins cysteine switch.
amino acids 85-92

Matrixin.
amino acids 38-199

FIGURE 115

MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTSVGAQNTVICSKLAAKCLVMKAEMNGSK
LGRRAKPEGALQNNDGLYDPDCDESGLFKAKQCNGTSTCWCVNTAGVRRTDKDTEITCSERVRTYWIIIELKHKARE
KPYDSKSLRTALQKEITTRYQLDPKFITSILYENNVITIDLVQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSKKM
DLTVNGEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVVMAVVAGIVVLVISRKKRMAKYEKAEIKEMGEM
HRELNA

Signal sequence.
amino acids 1-23

Transmembrane domain.
amino acids 265-285

N-glycosylation sites.
amino acids 74-77, 111-114, 198-201

Tyrosine kinase phosphorylation site.
amino acids 151-157

N-myristoylation sites.
amino acids 10-15, 52-57, 75-80, 86-91, 103-108, 112-117, 123-128, 263-268

Amidation site.
amino acids 78-81

Thyroglobulin type-1 repeat.
amino acids 66-135

FIGURE 116

MRQKAVSVFLCYLLLFTCSGVEAGKKKCSESSDSGSGFWKALTFMAVGGGLAVAGLPALGFTGAGIAANSVAASLMS
WSAILNGGGVPAGGLVATLQSLGAGGSSVVIGNIGALMRYATHKYLDSEEDEE

Signal sequence.
amino acids 1-20

Transmembrane domains.
amino acids 39-59, 90-110

Glycosaminoglycan attachment site.
amino acids 34-37

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 26-29

N-myristoylation sites.
amino acids 20-25, 48-53, 50-55, 63-68, 65-70, 85-90, 86-91, 90-95, 91-96, 100-105, 109-114

Amidation site.
amino acids 23-26

FIGURE 117

MPRAPAPLYACLLGLCALLPRLAGLNICTSGSATSCEECLLIHPKCAWCSKEDFGSPRSITSRCDLRANLVKNGCGG
EIESPASSFHVLRSLPLSSKGSGSAGWDVIQMTPQEIAVNLRPGDKTTFQLQVRQVEDYPVDLYYLMDLSLSMKDDL
DNIRSLGTKLAEEMRKLTSNFRLGFGSFVDKDISPFSYTAPRYQTNPCIGYKLFPNCVPSFGFRHLLPLTDRVDSFN
EEVRKQRVSRNRDAPEGGFDAVLQAAVCKEKIGWRKDALHLLVFTTDDVPHIALDGKLGGLVQPHDGQCHLNEANEY
TASNQMDYPSLALLGEKLAENNINLIFAVTKNHYMLYKNFTALIPGTTVEILDGDSKNIIQLIINAYNSIRSKVELS
VWDQPEDLNLFFTATCQDGVSYPGQRKCEGLKIGDTASFEVSLEARSCPSRHTEHVFALRPVGFRDSLEVGVTYNCT
CGCSVGLEPNSARCNGSGTYVCGLCECSPGYLGTRCECQDGENQSVYQNLCREAEGKPLCSGRGDCSCNQCSCFESE
FGKIYGPFCECDNFSCARNKGVLCSGHGECHCGECKCHAGYIGDNCNCSTDISTCRGRDGQICSERGHCLCGQCQCT
EPGAFGEMCEKCPTCPDACSTKRDCVECLLLHSGKPDNQTCHSLCRDEVITWVDTIVKDDQEAVLCFYKTAKDCVMM
FTYVELPSGKSNLTVLREPECGNTPNAMTILLAVVGSILLVGLALLAIWKLLVTIHDRREFAKFQSERSRARYEMAS
NPLYRKPISTHTVDFTFNKFNKSYNGTVD

Signal sequence.
amino acids 1-24
Transmembrane domain.
amino acids 723-743
N-glycosylation sites.
amino acids 347-350, 460-463, 477-480, 505-508, 552-555, 586-589, 654-657, 705-
708, 791-794, 795-798
Glycosaminoglycan attachment sites.
amino acids 523-526, 564-567
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 169-172
N-myristoylation sites.
amino acids 24-29, 31-36, 55-60, 98-103, 161-166, 248-253, 409-414, 415-420,
456-461, 464-469, 468-473, 480-485, 485-490, 560-565, 599-604, 611-616, 715-720
Cell attachment sequence.
amino acids 525-527
MotEGF-like domain cysteine pattern signatures.
amino acids 487-498, 574-585
Integrins beta chain cysteine-rich domain signatures.
amino acids 522-535, 563-576

FIGURE 118

MLPQIPFLLLVSLNLVHGVFYAERYQMPTGIKGPLPNTKTQFFIPYTIKSKGIAVRGEQGTPGPPGPAGPRGHPGPS
GPPGKPGYGSPGLQGEPGLPGPPGPSAVGKPGVPGLPGKPGERGPYGPKGDVGPAGLPGPRGPPGPPGIPGPAGISV
PGKPGQQGPTGAPGPRGFPGEKGAPGVPGMNGQKGEMGYGAPGRPGERGLPGPQGPTGPSGPPGVGKRGENGVPGQP
GIKGDRGFPGEMGPIGPPGPQGPPGERGPEGIGKPGAAGAPGQPGIPGTKGLPGAPGIAGPPGPPGFGKPGLPGLKG
ERGPAGLPGGPGAKGEQGPAGLPGKPGLTGPPGNMGPQGPKGIPGSHGLPGPKGETGPAGPAGYPGAKGERGSPGSD
GKPGYPGKPGLDGPKGNPGLPGPKGDPGVGGPPGLPGPVGPAGAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIPGF
PGSKGDPGSPGPPGPAGIATKGLNGPTGPPGPPGPRGHSGEPGLPGPPGPPGPPGQAVMPEGFIKAGQRPSLSGTPL
VSANQGVTGMPVSAFTVILSKAYPAIGTPIPFDKILYNRQQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHVWVGLYK
NGTPVMYTYDEYTKGYLDQASGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSSFSGFLVAPM

Signal sequence.
amino acids 1-18

Tyrosine kinase phosphorylation site.
amino acids 116-123

N-myristoylation sites.
amino acids 18-23, 276-281, 317-322, 350-355, 380-385, 415-420, 446-451,
529-534, 548-553, 587-592, 613-618, 661-666

Amidation site.
amino acids 219-224

C1q domain signature.
amino acids 571-601

C1q domain.
amino acids 553-677

Collagen triple helix repeat (20 copies).
amino acids 92-150, 156-214, 223-281, 282-340, 344-403, 404-463, 464-522

FIGURE 119

MLLWSLLVIFDAVTEQADSLTLVAPSSVFEGDSIVLKCQGEQNWKIQKMAYHKDNKELSVFKKFSDFLIQSAVLSDS
GNYFCSTKGQLFLWDKTSNIVKIKVQELFQRPVLTASSFQPIEGGPVSLKCETRLSPQRLDVQLQFCFFRENQVLGS
GWSSSPELQISAVWSEDTGSYWCKAETVTHRIRKQSLQSQIHVQRIPISNVSLEIRAPGGQVTEGQKLILLCSVAGG
TGNVTFSWYREATGTSMGKKTQRSLSAELEIPAVKESDAGKYYCRADNGHVPIQSKVVNIPVRIPVSRPVLTLRSPG
AQAAVGDLLELHCEALRGSPPILYQFYHEDVTLGNSSAPSGGGASFNLSLTAEHSGNYSCEANNGLGAQCSEAVPVS
ISGPDGYRRDLMTAGVLWGLFGVLGFTGVALLLYALFHKISGESSATNEPRGASRPNPQEFTYSSPTPDMEELQPVY
VNVGSVDVDVVYSQVWSMQQPESSANIRTLLENKDSQVIYSSVKKS

Signal sequence.
amino acids 1-14

Transmembrane domain.
amino acids 400-420

N-glycosylation sites.
amino acids 204-207, 234-237, 343-346, 355-358, 365-368

Glycosaminoglycan attachment site.
amino acids 348-351

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 62-65, 187-190

Tyrosine kinase phosphorylation site.
amino acids 266-273

N-myristoylation sites.
amino acids 78-83, 121-126, 153-158, 173-178, 213-218, 230-235, 245-250, 308-313, 349-354, 351-356, 364-369, 375-380, 400-405

Amidation site.
amino acids 248-251

Immunoglobulin domains.
amino acids 17-84, 121-179, 219-277, 314-370

FIGURE 120

LRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLYHLGRTVLCIDFMVFTVR
LLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQED
MDVALMEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQGNSDLYWK
AQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLAR
ARDKRESDSERLKRTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD

Transmembrane domains.
amino acids 34-54, 63-83, 99-119, 189-209, 238-258

N-glycosylation site.
amino acids 164-167

Glycosaminoglycan attachment site.
amino acids 6-9

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 312-315, 321-324

N-myristoylation sites.
amino acids 4-9, 7-12, 9-14, 10-15, 177-182, 181-186

Leucine zipper pattern.
amino acids 41-62

FIGURE 121

MGPSCPVFLSFTKLSLWWLLLTPAGGEEAKRPPPRAPGDPLSSPSPTALPQGGSHTETEDRLFKHLFRGYNRWARPV
PNTSDVVIVRFGLSIAQLIDVDEKNQMMTTNVWLKQEWSDYKLRWNPADFGNITSLRVPSEMIWIPDIVLYNNADGE
FAVTHMTKAHLFSTGTVHWVPPAIYKSSCSIDVTFFPFDQQNCKMKFGSWTYDKAKIDLEQMEQTVDLKDYWESGEW
AIVNATGTYNSKKYDCCAEIYPDVTYAFVIRRLPLFYTINLIIPCLLISCLTVLVFYLPSDCGEKITLCISVLLSLT
VFLLLITEIIPSTSLVIPLIGEYLLFTMIFVTLSIVITVFVLNVHHRSPSTHTMPHWVRGALLGCVPRWLLMNRPPP
PVELCHPLRLKLSPSYHWLESNVDAEEREVVVEEEDRWACAGHVAPSVGTLCSHGHLHSGASGPKAEALLQEGELLL
SPHMQKALEGVHYIADHLRSEDADSSVKEDWKYVAMVIDRIFLWLFIIVCFLGTIGLFLPPFLAGMI

Signal sequence.
amino acids 1-26

Transmembrane domains.
amino acids 264-284, 299-319, 330-350, 501-521

N-glycosylation sites.
amino acids 79-82, 129-132, 235-238

Tyrosine kinase phosphorylation site.
amino acids 468-475

N-myristoylation sites.
amino acids 25-30, 52-57, 89-94, 128-133, 238-243, 368-373, 434-439

Neurotransmitter-gated ion-channel signature.
amino acids 183-197

Neurotransmitter-gated ion-channel ligand binding domain.
amino acids 59-265

Neurotransmitter-gated ion-channel transmembrane region.
amino acids 272-520

FIGURE 122

MESRKDITNQEELWKMKPRRNLEEDDYLHKDTGETSMLKRPVLLHLHQTAHADEFDCPSELQHTQELFPQWHLPIKI
AAIIASLTFLYTLLREVIHPLATSHQQYFYKIPILVINKVLPMVSITLLALVYLPGVIAAIVQLHNGTKYKKFPHWL
DKWMLTRKQFGLLSFFFAVLHAIYSLSYPMRRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEIYVSLGIVGLAILA
LLAVTSIPSVSDSLTWREFHYIQSKLGIVSLLLGTIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLIFKSIL
FLPCLRKKILKIRHGWEDVTKINKTEICSQL

Transmembrane domains.
amino acids 72-92, 116-136, 158-178, 218-238, 254-274, 291-311

N-glycosylation sites.
amino acids 143-146, 331-334

Tyrosine kinase phosphorylation site.
amino acids 19-27

N-myristoylation sites.
amino acids 133-138, 265-270

FIGURE 123

MHNTTEKPTDAYGELDFTGAGRKHSNFLRLSDRTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRR
GLVRAAQSTGAWIVTGGLHTGIGRHVGVAVRDHQMASTGGTKVVAMGVAPWGVVRNRDTLINPKGSFPARYRWRGDP
EDGVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQQKTGVGGTGIDIPVLLLLIDGDEKMLTRIENATQA
QLPCLLVAGSGGAADCLAETLEDTLAPGSGGARQGEARDRIRRFFPKGDLEVLQAQVERIMTRKELLTVYSSEDGSE
EFETIVLKALVKACGSSEASAYLDELRLAVAWNRVDIAQSELFRGDIQWRSFHLEASLMDALLNDRPEFVRLLISHG
LSLGHFLTPMRLAQLYSAAPSNSLIRNLLDQASHSAGTKAPALKGGAAELRPPDVGHVLRMLLGKMCAPRYPSGGAW
DPHPGQGFGESMYLLSDKATSPLSLDAGLGQAPWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARL
EPDAEEAARRKDLAFKFEGMGVDLFGECYRSSEVRAARLLLRRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQK
WWGDMASTTPIWALVLAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPLGVPRQSGRP
GCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGSLELLLYFWAFTLLCEELRQGL
SGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLYHLGRTVLCIDFMVFTVRLLHIF
TVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVAL
MEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQGNSDLYWKAQRYR
LIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLARARDKR
ESDSERLERTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVT

Transmembrane domains.
amino acids 621-641, 713-733, 798-818, 828-848, 864-884, 954-974, 1003-1023
N-glycosylation sites.
amino acids 3-6, 165-168, 227-230, 929-932
Glycosaminoglycan attachment site.
amino acids 771-774
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 22-25, 1077-1080
N-myristoylation sites.
amino acids 78-83, 93-98, 94-99, 129-134, 142-147, 175-180, 178-183, 200-205,
240-245, 243-248, 259-264, 262-267, 323-328, 385-390, 467-472, 694-699, 697-702,
701-706, 714-719, 722-727, 769-774, 772-777, 774-779, 775-780, 942-947, 946-951
Amidation sites.
amino acids 20-23, 701-704
Cell attachment sequences.
amino acids 151-153, 352-354
Leucine zipper pattern.
amino acids 806-827

FIGURE 124

```
MHLHRRFTDLIWKNLCPALIVILGNPIHDKTITSAHTSSTSTSLESDSASPGVSDHGRGSGCSCTAPALSGPVARTI
YYIAAELVRLVGSVDSMKPVLQSLYHRVLLYPPPQHRVEAIKIMKEILGSPQRLCDLAGPSSTESESRKRSISKRKS
HLDLLKLIMDGMTEACIKGGIEACYAAVSCVCTLLGALDELSQGKGLSEGQVQLLLLRLEELKDGAEWSRDSMEINE
ADFRWQRRVLSSEHTPWESGNERSLDISISVTTDTGQTTLEGELGQTTPEDHSGNHKNSLKSPAIPEGKETLSKVLE
TEAVDQPDVVQRSHTVPYPDITNFLSVDCRTRSYGSRYSESNFSVDDQDLSRTEFDSCDQYSMAAEKDSGRSDVSDI
GSDNCSLADEEQTPRDCLGHRSLRTAALSLKLLKNQEADQHSARLFIQSLEGLLPRLLSLSNVEEVDTALQNFASTF
CSGMMHSPGFDGNSSLSFQMLMNADSLYTAAHCALLLNLKLSHGDYYRKRPTLAPGVMKDFMKQVQTSGVLMVFSQA
WIEELYHQVLDRNMLGEAGYWGSPEDNSLPLITMLTDIDGLESSAIGGQLMASAATESPFAQSRRIDDSTVAGVAFA
RYILVGCWKNLIDTLSTPLTGRMAGSSKGLAFILGAEGIKEQNQKERDAICMSLDGLRKAARLSCALGVAANCASAL
AQMAAASCVQEEKEEREAQEPSDAITQVKLKVEQKLEQIGKVQGVWLHTAHVLCMEAILSVGLEMGSHNPDCWPHVF
RVCEYVGTLEHNHFSDGASQPPLTISQPQKATGSAGLLGDPECEGSPPEHSPEQGRSLSTAPVVQPLSIQDLVREGS
RGRASDFRGGSLMSGSSAAKVVLTLSTQADRLFEDATDKLNLMALGGFLYQLKKASQSQLFHSVTDTVDYSLAMPGE
VKSTQDRKSALHLFRLGNAMLRIVRSKARPLLHVMRCWSLVAPHLVEAACHKERHVSQKAVSFIHDILTEVLTDWNE
PPHFHFNEALFRPFERIMQLELCDEDVQDQVVTSIGELVEVCSTQIQSGWRPLFSALETVHGGNKSEMKEYLVGDYS
MGKGQAPVFDVFEAFLNTDNIQVFANAATSYIMCLMKFVKGLGEVDCKEIGDCAPAPGAPSTDLCLPALDYLRRCSQ
LLAKIYKMPLKPIFLSGRLAGLPRRLQEQSASSEDGIESVLSDFDDDTGLIEVWIILLEQLTAAVSNCPRQHQPPTL
DLLFELLRDVTKTPGPGFGIYAVVHLLLPVMSVWLRRSHKDHSYWDMASANFKHAIGLSCELVVEHIQSFLHSDIRY
ESMINTMLKDLFELLVACVAKPTETISRVGCSCIRYVLVTAGPVFTEEMWRLACCALQDAFSATLKPVKDLLGCFHS
GTESFSGEGCQVRVAAPSSSPSAEAEYWRIRAMAQQVFMLDTQCSPKTPNNFDHAQSCQLIIELPPDEKPNGHTKKS
VSFREIVVSLLSHQVLLQNLYDILLEEFVKGPSPGEEKTIQVPEAKLAGFLRYISMQNLAVIFDLLLDSYRTAREFD
TSPGLKCLLKKVSGIGGAANLYRQSAMSFNIYFHALVCAVLTNQETITAEQVKKVLFEDDERSTDSSQQCSSEDEDI
FEETAQVSPPRGKEKRQWRARMPLLSVQPVSNADWVWLVKRLHKLCMELCNNYIQMHLDLENCMEEPPIFKGDPFFI
LPSFQSESSTPSTGGFSGKETPSEDDRSQSREHMGESLSLKAGGGDLLLPPSPKVEKKDPSRKKEWWENAGNKIYTM
AADKTISKLMTEYKKRKQQHNLSAFPKEVKVEKKGEPLGPRGQDSPLLQRPQHLMDQGQMRHSFSAGPELLRQDKRP
RSGSTGSSLSVSVRDAEAQIQAWTNMVLTVLNQIQILPDQTFTALQPAVFPCISQLTCHVTDIRVRQAVREWLGRVG
RVYDIIV
```

Transmembrane domains.

amino acids 482-502, 1243-1263, 1510-1530

N-glycosylation sites.

amino acids 350-353, 389-392, 475-478, 1065-1068, 1792-1795

Glycosaminoglycan attachment sites.

amino acids 1392-1395, 1553-1556

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 5-8, 145-148, 151-154, 511-514, 900-903, 1151-1154, 1550-1153

Tyrosine kinase phosphorylation site.

amino acids 172-179, 1269-1276

N-myristoylation sites.

amino acids 59-64, 61-66, 89-94, 165-170, 173-178, 174-179, 200-205, 219-224,
251-256, 285-290, 343-348, 386-391, 579-584, 587-592, 612-617, 622-627, 641-646,
672-677, 684-689, 755-760, 857-862, 862-867, 1063-1068, 1136-1141, 1382-1387,
1554-1559, 1556-1561, 1708-1713, 1851-1856, 1854-1859

FIGURE 125

MVDPNGNESSATYFILIGLPGLEEAQFWLAFPLCSLYLIAVLGNLTIIYIVRTEHSLHEPMYIFLCMLSGIDILIST
SSMPKMLAIFWFNSTTIQFDACLLQIFAIHSLSGMESTVLLAMAFDRYVAICHPLRHATVLTLPRVTKIGVAAVVRG
AALMAPLPVFIKQLPFCRSNILSHSYCLHQDVMKLACDDIRVNVVYGLIVIISAIGLDSLLISFSYLLILKTVLGLT
REAQAKAFGTCVSHVCAVFIFYVPFIGLSMVHRFSKRRDSPLPVILANIYLLVPPVLNPIVYGVKTKEIRQRILRLF
HVATHASEP

Transmembrane domains.
amino acids 27-47, 61-81, 99-119, 141-161, 205-225, 239-259, 273-293

N-glycosylation sites.
amino acids 7-10, 44-47, 90-93

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 268-271

N-myristoylation sites.
amino acids 6-11, 21-26, 111-116, 240-245

7 transmembrane receptor (rhodopsin family).
amino acids 43-293

FIGURE 126

MGSNSGQAGRHIYKSLADDGPFDSVEPPKRPTSRLIMHSMAMFGREFCYAVEAAYVTPVLLSVGLPSSLYSIVWFLS
PILGFLLQPVVGSASDHCRSRWGRRRPYILTLGVMMLVGMALYLNGATVVAALIANPRRKLVWAISVTMIGVVLFDF
AADFIDGPIKAYLFDVCSHQDKEKGLHYHALFTGFGGALGYLLGAIDWAHLELGRLLGTEFQVMFFFSALVLTLCFT
VHLCSISEAPLTEVAKGIPPQQTPQDPPLSSDGMYEYGSIEKVKNGYVNPELAMQGAKNKNHAEQTRRAMTLKSLLR
ALVNMPPHYRYLCISHLIGWTAFLSNMLFFTDFMGQIVYRGDPYSAHNSTEFLIYERGVEVGCWGFCINSVFSSLYS
YFQKVLVSYIGLKGLYFTGYLLFGLGTGFIGLFPNVYSTLVLCSLFGVMSSTLYTVPFNLITEYHREEEKERQQAPG
GDPDNSVRGKGMDCATLTCMVQLAQILVGGGLGFLVNTAGTVVVVVITASAVALIGCCFVALFVRYVD

Transmembrane domains.
amino acids 68-88, 105-125, 141-161, 182-202, 216-236, 318-338, 394-414, 422-
442, 501-521

N-glycosylation site.
amino acids 356-359

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 29-32

N-myristoylation sites.
amino acids 2-7, 64-69, 190-195, 366-371, 399-404, 409-414, 416-421, 432-437,
473-478, 491-496

Amidation site.
amino acids 99-102

Cell attachment sequence.
amino acids 348-350

FIGURE 127

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQLQSSDFHSVSKLRVL
ILCHNNRIQQLDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDLSFNDFDTMPICEEAGNMSHLEILGLSGAKI
QKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLPILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSY
EMQRNLSLENAKTSVLLLNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVH
FRVFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLPHLKTLILNGNKLE
TLSLVSCFANNTPLEHLDLSQNLLQHKNDENCSWPETVVNMNLSYNKLSDSVFRCLPKSIQILDLNNNQIQTVPKET
IHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIEMNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETY
SEVMMVGWSDSYTCEYPLNLRGTRLKDVHLHELSCNTALLIVTIVVIMLVLGLAVAFCCLHFDLPWYLRMLGQCTQT
WHRVRKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENIVSFIEKSYKSI
FVLSPNFVQNEWCHYEFYFAHHNLFHENSDHIILILLLEPIPFYCIPTRYHKLKALLEKKAYLEWPKDRRKCGLFWAN
LRAAINVNVLATREMYELQTFTELNEESRGSTISLMRTDCL

Transmembrane domain.
amino acids 577-597
N-glycosylation sites.
amino acids 33-36, 36-39, 140-143, 189-192, 236-239, 278-281, 330-333, 416-419,
427-430
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 621-624
Tyrosine kinase phosphorylation site.
amino acids 746-754
N-myristoylation sites.
amino acids 148-153, 210-215, 591-596, 765-770, 800-805
Leucine zipper patterns.
amino acids 39-60, 101-122
Nt-dnaJ domain signature.
amino acids 350-369
TIR domain.
amino acids 636-774
Leucine rich repeats.
amino acids 49-72, 73-96, 97-120, 143-166, 373-394, 398-418, 422-442, 444-466,
467-488, 489-512
Leucine rich repeat C-terminal domain.
amino acids 522-575

FIGURE 128

MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPPWGKYDAALRTMIPFRPK
PRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPLSVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLV
MLRFQRTRLIFDALLGICFCIASVLGPILIIPKILEYSEEQLGNVVHGVGLCFALFLSECVKSLSFSSSWIINQRTA
IRFRAAVSSFAFEKLIQFKSVIHITSGEAISFFTGDVNYLFEGVCYGPLVLITCASLVICSISSYFIIGYTAFIAIL
CYPLVFPLEVFMTRMAVKAQHHTSEVSDQRIRVTSEVLTCIKLIKMYTWEKPFAKIIEDLRRKERKLLEKCGLVQSL
TSITLFIIPAVATAVWVLIHTSLKLKLTASMAFSMLASLNLLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVF
YVQTLQDPSKALVFEEATLSWQQTCPGIVNGALELERNGHASEGMTRPRDALGPEEEGNSLGPELHKINLVVSKGMM
LGVCGNTGSGKSSLLSAILEEMHLLEGSVGVQGSLAYVPQQAWIVSGNIRENILMGGAYDKARYLQVLHCCSLNRDL
ELLPFGDMTEIGERGPNLSGGQKQRISLARAVYSDRQIYLLDDPLSAVDAHVGKHIFEECIKKTLRGKTVVQVTHQL
QYLEFCGQVILLENGKICENGTHSELMQKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLEESLNGN
AVPEHQLTQEEEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFFVVLIVFLTIFSFWWLSYWLEQGSGTNSSRESNGTM
ADLGNIADNPQLSFYQLVYGLNALLLICVGVCSSGIFTKVTRKASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCFA
GDLEQLDQLLPIFSEQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIIYMMFKEAIGVFKRLENYSRSPLFSH
ILNSLQGLSSIHVYGKTEDFISQFKRLTDAQNNYLLLFLSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSFKV
MAVNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPQGWPQHGEIIFQDYHMKYRDNTP
TVLHGINLTIRGHEVVGIVGRTGSGKSSLGMALFRLVEPMAGRILIDGVDICSIGLEDLRSKLSVIPQDPVLLSGTI
RFNLDPFDRHTDQQIWDALERTFLTKAISKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASID
METDTLIQRTIREAFQGCTVLVIAHRVTTVLNCDRILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR

Transmembrane domains.

amino acids 163-183, 199-219, 270-290, 300-320, 381-401, 418-438, 804-824, 858-
878, 902-922, 935-955, 951-971, 990-1010, 1042-1062

N-glycosylation sites.

amino acids 11-14, 633-636, 713-716, 838-841, 844-847, 992-925, 1162-1165, 1277-
1280

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 3-6, 889-892, 1026-1029

N-myristoylation sites.

amino acids 14-19, 20-25, 145-150, 170-175, 202-207, 204-209, 380-385, 440-445,
489-494, 537-542, 541-546, 544-549, 547-552, 569-574, 834-839, 836-841, 877-
882, 1160-1165, 1178-1183, 1203-1208, 1275-1280, 1367-1372

ATP/GTP-binding site motif A (P-loop).

amino acids 544-551, 1175-1182

ABC transporter.

amino acids 537-708, 1168-1351

ABC transporter transmembrane region.

amino acids 163-431, 806-1095

FIGURE 129

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAVQLSPAIPVFAAMLFL
FSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPPRIKNFQINNQIVKLKYCYTCKIFRPPRASH
CSICDNCVERFDHHCPWVGNCVGKRNYRYFYLFILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLI
CFFTLWSVVGLTGFHTFLVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESG
SRPPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

Transmembrane domains.
amino acids 32-52, 66-86, 184-204, 227-247

N-glycosylation site.
amino acids 253-256

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 8-11

N-myristoylation sites.
amino acids 173-178, 262-267

Amidation site.
amino acids 176-179

DHHC zinc finger domain.
amino acids 130-194

FIGURE 130

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALVGLIGAVKHHQVLLFF
YMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGWNNTASARNDIQRNLNCCGFRSVNPNDTCLASCVKSDHSCSPC
APIIGEYAGEVLRFVGGIGLFFSFTEILGVWLTYRYRNQKDPRANPSAFL

Transmembrane domains.
amino acids 10-30, 43-63, 75-95, 163-183

N-glycosylation sites.
amino acids 113-116, 137-140

N-myristoylation sites.
amino acids 4-9, 34-39, 38-43, 47-52, 63-68, 173-178, 183-188

Tetraspanin family homology.
amino acids 10-194

FIGURE 131

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQVKAYTFSEPFHLIVSY
DWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGPNREFSITVVQKADSGHYHCSGIFQSPGPGI
PETASVVAITVQELFPAPILRAVPSAEPQAGSPMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTAS
EDHSGSYWCEAATEDNQVWKQSPQLEIRVQGASSSAAPPTLNPAPQKSAAPGTAPEEAPGPLPPPPTPSSEDPGFSS
PLGMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence.
amino acids 1-17

N-myristoylation sites.
amino acids 153-158, 185-190, 236-241, 262-267, 352-357

Leucine zipper pattern.
amino acids 12-33

FIGURE 132

MPGGCSRGPAAGDGRLRLARLALVLLGWVSSSSPTSSASSFSSSAPFLASAVSAQPPLPDQCPALCECSEAARTVKC
VNRNLTEVPTDLPAYVRNLFLTGNQLAVLPAGAFARRPPLAELAALNLSGSRLDEVRAGAFEHLPSLRQLDLSHNPL
ADLSPFAFSGSNASVSAPSPLVELILNHIVPPEDERQNRSFEGMVVAALLAGRALQGLRRLELASNHFLYLPRDVLA
QLPSLRHLDLSNNSLVSLTYVSFRNLTHLESLHLEDNALKVLHNGTLAELQGLPHIRVFLDNNPWVCDCHMADMVTW
LKETEVVQGKDRLTCAYPEKMRNRVLLELNSADLDCDPILPPSLQTSYVFLGIVLALIGAIFLLVLYLNRKGIKKWM
HNIRDACRDHMEGYHYRYEINADPRLTNLSSNSDV

Signal sequence.
amino acids 1-31

Transmembrane domain.
amino acids 355-375

N-glycosylation sites.
amino acids 81-84, 124-127, 166-169, 192-195, 243-246, 256-259, 275-278, 413-416

Tyrosine kinase phosphorylation site.
amino acids 393-401

N-myristoylation sites.
amino acids 100-105, 164-169, 197-199, 360-365

Leucine zipper pattern.
amino acids 82-103

Leucine rich repeats.
amino acids 92-115, 119-142, 143-166, 211-234, 235-258, 259-282

Leucine rich repeat C-terminal domain.
amino acids 294-345

Leucine rich repeat N-terminal domain.
amino acids 61-90

FIGURE 133

MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTSVGAQNTVICSKLAAKCLVMKAEMNGSK
LGRRAKPEGALQNNDGLYDPDCDESGLFKAKQCNGTSTCWCVNTAGVRRTDKDTEITCSERVRTYWIIIELKHKARE
KPYDSKSLRTALQKEITTRYQLDPKFITSILYENNVITIDLVQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSKKM
DLTVNGEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMAKYEKAEIKEMGEM
HRELNA

Signal sequence.
amino acids 1-23

Transmembrane domain.
amino acids 265-285

N-glycosylation sites.
amino acids 74-77, 111-114, 198-201

Tyrosine kinase phosphorylation site.
amino acids 151-157

N-myristoylation sites.
amino acids 10-15, 52-57, 75-80, 86-91, 103-108, 112-117, 123-128, 263-268

Amidation site.
amino acids 78-81

Thyroglobulin type-1 repeat.
amino acids 66-135

FIGURE 134A

MPCGFSPSPVAHHLVPGPPDTPAQQLRCGWTVGGWLLSLVRGLLPCLPPGARTAEGPIMVLAGPLAVSLLLPSLTLL
VSHLSSQDVSSEPSSEQQLCALSKHPTVAFEDLQPWVSNFTYPGARDFSQLALDPSGNQLIVGARNYLFRLSLANV
SLLQATEWASSEDTRRSCQSKGKTEEECQNYVRVLIVAGRKVFMCGTNAFSPMCTSRQVGNLSRTIEKINGVARCPY
DPRHNSTAVISSQGELYAATVIDFSGRDPAIYRSLGSGPPLRTAQYNSKWLNEPNFVAAYDIGLFAYFFLRENAVEH
DCGRTVYSRVARVCKNDVGGRFLLEDTWTTFMKARLNCSRPGEVPFYYNELQSAFHLPEQDLIYGVFTTNVNSIAAS
AVCAFNLSAISQAFNGPFRYQENPRAAWLPIANPIPNFQCGTLPETGPNENLTERSLQDAQRLFLMSEAVQPVTPEP
CVTQDSVRFSHLVVDLVQAKDTLYHVLYIGTESGTILKALSTASRSLHGCYLEELHVLPPGRREPLRSLRILHSARA
LFVGLRDGVLRVPLERCAAYRSQGACLGARDPYCGWDGKQQRCSTLEDSSNMSLWTQNITACPVRNVTRDGGFGPWS
PWQPCEHLDGDNSGSCLCRARSCDSPRPRCGGLDCLGPAIHIANCSRNGAWTPWSSWALCSTSCGIGFQVRQRSCSN
PAPRHGGRIFVGKSREERFCNENTPCPVPIFWASWGSWSKCSSNCGGGMQSRRRACENGNSCLGCGEFKTCNPEGCP
EVRRNTPWTPWLPVNVTQGGARQEQRFRFTCRAPLADPHGLQFGRRRTETRTCPADGSGSCDTDALVEVLLRSGSTS
PHTVSGGWAAWGPWSSCSRDCELGFRVRKRTCTNPEPRNGGLPCVGDAAEYQDCNPQACPVRGAWSCWTSWSPCSAS
CGGGHYQRTRSCTSPAPSPGEDICLGLHTEEALCATQACPGWSPWSEWSKCTDDGAQSRSRHCEELLPGSSACAGNS
SQSRPCPYSEIPVILPASSMEEATDCAGKRNRTYLMLRSSQPSSTPLQSLDSFHILLQTAKLCWGPHCFEMGSISST
WWPRASPASWALGS

Signal sequence.

amino acids 1-42

Transmembrane domain.

amino acids 60-80

N-glycosylation sites.

amino acids 117-120, 153-156, 215-218, 236-239, 345-348, 391-394, 436-439, 590-
593, 597-600, 605-608, 660-663, 785-788, 1000-1003, 1032-1035

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 773-776, 815-818, 875-878

Tyrosine kinase phosphorylation sites.

amino acids 177-185, 348-356

N-myristoylation sites.

amino acids 42-47, 50-55, 373-378, 492-497, 543-548, 563-568, 630-635, 647-652,
740-745, 810-815, 827-832, 829-834, 853-858, 887-892, 910-915, 993-998, 1073-
1078

Amidation sites.

amino acids 192-195, 522-525, 813-816, 1028-1031

ATP/GTP-binding site motif A (P-loop).

amino acids 700-707

Cytochrome c oxidase subunit II, copper A binding region signature.

amino acids 921-929

Growth factor and cytokines receptors family signature 2.

amino acids 967-973

Sema domain.

amino acids 126-537

FIGURE 134B

Thrombospondin type 1 domains.
amino acids 613-661, 668-719, 726-769, 856-906, 913-963, 967-1007
Plexin repeat.
amino acids 555-602
Plant PEC family metallothionein.
amino acids 712-791

FIGURE 135

MAKDNSTVRCFQGLLIFGNVIIGCCGIALTAECIFFVSDQHSLYPLLEATDNDDIYGAAWIGIFVGICLFCLSVLGI
VGIMKSSRKILLAYFILMFIVYAFEVASCITAATQQDFFTPNLFLKQMLERYQNNSPPNNDDQWKNNGVTKTWDRLM
LQDNCCGVNGPSDWQKYTSAFRTENNDADYPWPRQCCVMNNLKEPLNLEACKLGVPGFYHNQGCYELISGPMNRHAW
GVAWFGFAILCWTFWVLLGTMFYWSRIEY

Transmembrane domains.
amino acids 12-32, 54-74, 89-109, 231-251

N-glycosylation site.
amino acids 5-8

N-myristoylation sites.
amino acids 26-31, 62-67, 79-84, 145-150

Tetraspanin family homology.
amino acids 12-258

FIGURE 136

MFDKTRLPYVALDVLCVLLAGLPFAIFTSRHITSRHTPFQRGVFCNDESIKYPYKEDTIPYALLGGIIIPFSIIVII
LGETLSVYCNLLHSNSFIRNNYIATIYKAIGTFLFGAAASQSLTDIAKYSIGRLRPHFLDVCDPDWSKINCSDGYIE
YYICRGNAERVKEGRLSFYSGHSSFSMYCMLFVALYLQARMKGDWARLLRPTLQFGLVAVSIYVGLSRVSDYKHHWS
DVLTGLIQGALVAILVAVYVSDFFKERTSFKERKEEDSHTTLHETPTTGNHYPSNHQP

Transmembrane domains.
amino acids 7-27, 61-81, 97-117, 172-192, 200-220, 233-253

N-glycosylation site.
amino acids 147-150

N-myristoylation sites.
amino acids 21-26, 42-47, 113-118, 236-241, 240-245

PAP2 superfamily.
amino acids 105-256

FIGURE 137

MLLWLLLLILTPGREQSGVAPKAVLLLNPPWSTAFKGEKVALICSSISHSLAQGDTYWYHDEKLLKIKHDKIQITEP
GNYQCKTRGSSLSDAVHVEFSPDWLILQALHPVFEGDNVILRCQGKDNKNTHQKVYYKDGKQLPNSYNLEKITVNSV
SRDNSKYHCTAYRKFYILDIEVTSKPLNIQVQELFLHPVLRASSSTPIEGSPMTLTCETQLSPQRPDVQLQFSLFRD
SQTLGLGWSRSPRLQIPAMWTEDSGSYWCEVETVTHSIKKRSLRSQIRVQRVPVSNVNLEIRPTGGQLIEGENMVLI
CSVAQGSGTVTFSWHKEGRVRSLGRKTQRSLLAELHVLTVKESDAGRYYCAADNVHSPILSTWIRVTVRIPVSHPVL
TFRAPRAHTVVGDLLELHCESLRGSPPILYRFYHEDVTLGNSSAPSGGGASFNLSLTAEHSGNYSCDADNGLGAQHS
HGVSLRVTVPVSRPVLTLRAPGAQAVVGDLLELHCESLRGSFPILYWFYHEDDTLGNISAHSGGGASFNLSLTTEHS
GNYSCEADNGLGAQHSKVVTLNVTGTSRNRTGLTAAGITGLVLSILVLAAAAALLHYARARRKPGGLSATGTSSHSP
SECQEPSSSRPSRIDPQEPTHSKPLAPMELEPMYSNVNPGDSNPIYSQIWSIQHTKENSANCPMMHQEHEELTVLYS
ELKKTHPDDSAGEASSRGRAHEEDDEENYENVPRVLLASDH

Signal sequence.
amino acids 1-13
Transmembrane domain.
amino acids 574-594
N-glycosylation sites.
amino acids 426-429, 438-441, 448-451, 519-522, 531-534, 541-544, 561-561, 568-571
Glycosaminoglycan attachment sites.
amino acids 431-434, 524-527
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 270-273
Tyrosine kinase phosphorylation site.
amino acids 349-356
N-myristoylation sites.
amino acids 78-83, 86-91, 204-209, 236-241, 256-261, 432-437, 434-439, 447-452, 458-463, 518-523, 525-530, 527-532, 540-545, 551-556, 564-569, 571-576, 579-584, 604-609, 605-610
Amidation site.
amino acids 331-334
N-6 Adenine-specific DNA methylases signature.
amino acids 25-31
Immunoglobulin domains.
amino acids 37-84, 113-165, 204-262, 302-360, 397-453, 490-546

FIGURE 138

MEGGAAAATPTALPYYVAFSQLLGLTLVAMTGAWLGLYRGGIAWESDLQFNAHPLCMVIGLIFLQGNALLVYRVFRN
EAKRTTKVLHGLLHIFALVIALVGLVAVFDYHRKKGYADLYSLHSWCGILVFVLYFVQWLVGFSFFLFPGASFSLRS
RYRPQHIFFGATIFLLPVGTALLGLKEALLFNLGGKYSAFEPEGVLANVLGLLLACFGGAVLYILTRADWKRPSQAE
EQALSMDFKTLRQGDSPGSQ

Transmembrane domains.
amino acids 13-33, 51-71, 85-105, 123-143, 161-181, 198-218

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 80-83, 225-228

N-myristoylation sites.
amino acids 3-8, 4-9, 32-37, 36-41, 147-152, 178-183, 188-193, 198-203, 205-210

Cytochrome b561 homology.
amino acids 1-238

FIGURE 139

MRQLCRGRVLGISVAIAHGVFSGSLNILLKFLISRYQFSFLTLVQCLTSSTAALSLELLRRLGLIAVPPFGLSLARS
FAGVAVLSTLQSSLTLWSLRGLSLPMYVVFKRCLPLVTMLIGVLVLKNGAPSPGVLAAVLITTCGAALAGAGDLTGD
PIGYVTGVLAVLVHAAYLVLIQKASADTEHGPLTAQYVIAVSATPLLVICSFASTDSIHAWTFPGWKDPAMVCIFVA
CILIGCAMNFTTLHCTYINSAVTTSLFIAGVVVNTLGSIIYCVAKFMETRKQSNYEDLEAQPRGEEAQLSGDQLPFV
MEELPGEGGNGRSEGGEAAGGPAQESRQEVRGSPRGVPLVAGSSEEGSRRSLKDAYLEVWRLVRGTRYMKKDYLIEN
EELPSP

Transmembrane domains.
amino acids 8-28, 51-71, 103-123, 119-139, 153-173, 189-209, 221-241, 253-273

N-glycosylation site.
amino acids 240-243

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 281-284

N-myristoylation sites.
amino acids 11-16, 19-24, 71-76, 131-136, 142-147, 236-241, 261-266, 317-322,
323-328, 340-345, 355-360

FIGURE 140

MARLALSPVPSHWMVALLLLLSAEPVPAARSEDRYRNPKGSACSRIWQSPRFIARKRGFTVKMHCYMNSASGNVSWL
WKQEMDENPQQLKLEKGRMEESQNESLATLTIQGIRFEDNGIYFCQQKCNNTSEVYQGCGTELRVMGFSTLAQLKQR
NTLKDGIIMIQTLLIILFIIVPIFLLLDKDDSKAGMEEDHTYEGLDIDQTATYEDIVTLRTGEVKWSVGEHPGQE

Signal sequence.
amino acids 1-28

Transmembrane domain.
amino acids 159-179

N-glycosylation sites.
amino acids 73-76, 101-104, 127-130, 128-131

Tyrosine kinase phosphorylation site.
amino acids 113-120

N-myristoylation sites.
amino acids 40-45, 118-123

Immunoglobulin domain.
amino acids 58-124

Immunoreceptor tyrosine-based activation motif.
amino acids 193-213

FIGURE 141

MGLPEPGPLRLLALLLLLLLLLLLRLQHLAAAAADPLLGGQGPAKECEKDQFQCRNERCIPSVWRCDEDDDCLDHSD
EDDCPKKTCADSDFTCDNGHCIHERWKCDGEEECPDGSDESEATCTKQVCPAEKLSCGPTSHKCVPASWRCDGEKDC
EGGADEAGCATSLGTCRGDEFQCGDGTCVLAIKHCNQEQDCPDGSDEAGCLQGLNECLHNNGGCSHICTDLKIGFEC
TCPAGFQLLDQKTCGDIDECKDPDACSQICVNYKGYFKCECYPGCEMDLLTKNCKAAAGKSPSLIFTNRTSAEDRPV
KRNYSRLIPMLKNVVALDVEVATNRIYWCDLSYRKIYSAYMDKASDPKEREVLIDEQLHSPEGLAVDWVHKHIYWTD
SGNKTISVATVDGGRRRTLFSRNLSEPRAIAVDPLRGFMYWSDWGDQAKIEKSGLNGVDRQTLVSDNIEWPNGITLD
LLSQRLYWVDSKLHQLSSIDFSGGNRKTLISSTDFLSHPFGIAVFEDKVFWTDLENEAIFSANRLNGLEISILAENL
NNPHDIVIFHELKQPRAPDACELSVQPNGGCEYLCLPAPQISSHSPKYTCACPDTMWLGPDMKRCYRDANEDSKMGS
TVTAAVIGIIVPIVVIALLCMSGYLIWRNWKRKNTKSMNFDNPVYRKTTEEEDEDELHIGRTAQIGHVYPARVALSL
EDDGLP

Signal sequence.
amino acids 1-26
Transmembrane domain.
amino acids 619-639
N-glycosylation sites.
amino acids 299-302, 311-314, 388-391, 408-411
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 400-403, 648-651, 662-665
N-myristoylation sites.
amino acids 40-45, 114-119, 157-162, 162-167, 168-173, 198-203, 203-208, 207-
212, 486-491, 529-534, 615-620
Amidation site.
amino acids 398-401
Aspartic acid and asparagine hydroxylation sites.
amino acids 222-233, 261-272
Cell attachment sequence.
amino acids 171-173
Low-density lipoprotein receptor repeats.
amino acids 332-377, 379-420, 422-464, 466-509, 510-550
Low-density lipoprotein receptor domains.
amino acids 45-83, 84-124, 125-165, 168-206

FIGURE 142

MPPLWALLALGCLRFGSAVNLQPQLASVTFATNNPTLTTVALEKPLCMFDSKEALTGTHEVYLYVLVDSAISRNASV
QDSTNTPLGSTFLQTEGGRTGPYKAVAFDLIPCSDLPSLDAIGDVSKASQILNAYLVRVGANGTCLWDPNFQGLCNA
PLSAATEYRFKYVLVNMSTGLVEDQTLWSDPIRTNQLTPYSTIDTWPGRRSGGMIVITSILGSLPFFLLVGFAGAIA
LSLVDMGSSDGETTHDSQITQEAVPKSLGASESSYTSVNRGPPLDRAEVYSSKLQD

Signal sequence.
amino acids 1-18

Transmembrane domain.
amino acids 211-231

N-glycosylation sites.
amino acids 74-77, 139-142, 170-173

N-myristoylation sites.
amino acids 16-21, 137-142, 238-243, 260-265

Amidation site.
amino acids 201-204

FIGURE 143

MRKLIAGLIFLKFWTYTVRASTDLPQTENCFQYIHQVTEISSTLPVALLRDEVPGWFLKVPEPQLISKELIMLTEVM
EVWHGLVIAVVSLFLQACFLTAINYLLSRHMAHKSEQILKAASLQVPRPSPGHHHPPAVKEMKETQTERDIPMSDSL
YRHDSDTPSDSLDSSCSSPPACQATEDVDYTQVVFSDPGELK

Signal sequence.
amino acids 1-20

Transmembrane domain.
amino acids 75-95

N-myristoylation site.
amino acids 82-87

Leucine zipper pattern.
amino acids 83-104

FIGURE 144

MAGTVRTACLVVAMLLSLDFPGQAQPPPPPPDATCHQVRSFFQRLQPGLKWVPETPVPGSDLQVCLPKGPTCCSRKM
EEKYQLTARLNMEQLLQSASMELKFLIIQNAAVFQEAFEIVVRHAKNYTNAMFKNNYPSLTPQAFEFVGEFFTDVSL
YILGSDINVDDMVNELFDSLFPVIYTQLMNPGLPDSALDINECLRGARRDLKVFGNFPKLIMTQVSKSLQVTRIFLQ
ALNLGIEVINTTDHLKFSKDCGRMLTRMWYCSYCQGLMMVKPCGGYCNVVMQGCMAGVVEIDKYWREYILSLEELVN
GMYRIYDMENVLLGLFSTIHDSIQYVQKNAGKLTTTIGKLCAHSQQRQYRSAYYPEDLFIDKKVLKVAHVEHEETLS
SRRRELIQKLKSFISFYSALPGYICSHSPVAENDTLCWNGQELMERYSQKAARNGMKNQFNLHELKMKGPEPVVSQI
IDKLKHINQLLRTMSMPKGRVLDKNLDEEGFESGDCGDDEDECIGGSGDGMIKVKNQLRFLAELAYDLDVDDAPGNS
QQATPKDNEISTFHNLGNVHSPLKLLTSMAISVVCFFFLVH

Signal sequence.
amino acids 1-24

Transmembrane domain.
amino acids 559-579

N-glycosylation sites.
amino acids 124-127, 241-244, 418-421

Glycosaminoglycan attachment site.
amino acids 509-512

Tyrosine kinase phosphorylation sites.
amino acids 75-81, 521-528

N-myristoylation sites.
amino acids 3-8, 158-163, 186-191, 275-280, 284-289, 322-327, 508-513

Glypican homology.
amino acids 4-578

FIGURE 145

MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEPGQALGKVFMGCPGQEPALFSTDNDDFTVRN
GETVQERRSLKERNPLKIFPSKRILRRHKRDWVVAPISVPENGKGPFPQRLNQLKSNKDRDTKIFYSITGPGADSPP
EGVFAVEKETGWLLLNKPLDREEIAKYELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTFRGSVLEGVLPG
TSVMQVTATDEDDAIYTNGVVAYSIHSQEPKDPHDLMFTIHRSTGTISVISSGLDREKVPEYTLTIQATDMDGDGS
TTTAVAVVEILDANDNAPMFDPQKYEAHVPENAVGHEVQRLTVTDLDAPNSPAWRATYLIMGGDDGDHFTITTHPES
NQGILTTRKGLDFEAKNQHTLYVEVTNEAPFVLKLPTSTATIVVHVEDVNEAPVFVPPSKVVEVQEGIPTGEPVCVY
TAEDPDKENQKISYRILRDPAGWLAMDPDSGQVTAVGTLDREDEQFVRNNIYEVMVLAMDNGSPPTTGTGTLLLTLI
DVNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPHTSPFQAQLTDDSDIYWTAEVNEEGDTVVLSLKKFLKQDTYD
VHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPWKGGFILPVLGAVLALLFLLLVLLLLVRKKRKIKEPLLLPEDD
TRDNVFYYGEEGGGEEDQDYDITQLHRGLEARPEVVLRNDVAPTIIPTPMYRPRPANPDEIGNFIIENLKAANTDPT
APPYDTLLVFDYEGSGSDAASLSSLTSSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD

Signal sequence.
amino acids 1-24

Transmembrane domain.
amino acids 655-675

N-glycosylation sites.
amino acids 200-203, 566-569

N-myristoylation sites.
amino acids 48-53, 60-65, 227-232, 307-312, 370-375, 388-393, 452-457, 493-498,
524-529, 661-666, 786-791

Cadherins extracellular repeated domain signature.
amino acids 203-213, 316-326, 536-546

Cadherin domains.
amino acids 122-206, 220-319, 333-432, 445-539, 551-638

Cadherin cytoplasmic region.
amino acids 678-826

FIGURE 146

GDCGDRGTARGTRREGTGIRSSGRAMDGNDNVTLLFAPLLRDNYTLAPNASSLGPGTDLALAPASSAGPGPGLSLGP
GPSFGFSPGPTPTPEPTTSGLAGGAASHGPSPFPRPWAPHALPFWDTPLNHGLNVFVGAALCITMLGLGCTVDVNHF
GAHVRRPVGALLAALCQFGLLPLLAFLLALAFKLDEVAAVAVLLCGCCPGGNLSNLMSLLVDGDMNLSIIMTISSTL
LALVLMPLCLWIYSWAWINTPIVQLLPLGTVTLTLCSTLIPIGLGVFIRYKYSRVADYIVKVSLWSLLVTLVVLFIM
TGTMLGPELLASIPAAVYVIAIFMPLAGYASGYGLATLFHLPPNCKRTVCLETGSQNVQLCTAILKLAFPPQFIGSM
YMFPLLYALFQSAEAGIFVLIYKMYGSEMLHKRDPLDEDEDTDISYKKLKEEEMADTSYGTVKAENIIMMETAQTSL

Transmembrane domains.
amino acids 126-146, 161-181, 177-197, 218-238, 254-274, 289-309, 326-346

N-glycosylation sites.
amino acids 31-34, 43-46, 49-52, 206-209, 220-223

Glycosaminoglycan attachment site.
amino acids 339-342

N-myristoylation sites.
amino acids 7-12, 18-23, 56-61, 97-102, 100-105, 135-140, 144-149, 163-168, 200-
205, 204-209, 205-210, 445-450

Sodium Bile acid symporter family homology.
amino acids 135-316

FIGURE 147

MVRRDRLRRMREWWVQVGLLAVPLLAAYLHIPPPQLSPALHSWKSSGKFFTYKGLRIFYQDSVGVVGSPEIVVLLHG
FPTSSYDWYKIWKGLTLRFHRVIALDFLGFGFSDKPRPHHYSIFEQASIVEALLRHLGLQNRRINLLSHDYGDIVAQ
ELLYRYKQNRSGRHTIKSLCLSNGGIFPETHRPLLLQKLLKDGGVLSPILTRLMNFFVFSRGLTPVFGPYTRPSESE
LWDMWAGIRNNDGNLVIDSLLQYINQRKKFRRRWVGALASVTIPIHFIYGPLDPVNPYPEFLELYRKTLPRSTVSIL
DDHISHYPQLEDPMGFLNAYMGFINSF

Transmembrane domains.
amino acids 12-32, 59-79

N-glycosylation site.
amino acids 163-166

N-myristoylation sites.
amino acids 238-243, 267-272

alpha/beta hydrolase fold.
amino acids 97-332

FIGURE 148

GWTSHLSNCGESNRPPKERSCFRVCDWHSDLFQWEVSDWHHCVLVPYARGEVKPRTAECVTAQHGLQHRMVRCIQKL
NRTVVANEICEHFALQPPTEQACLIPCPRDCVVSEFLPWSNCSKGCGKKLQHRTRAVIAPPLFGGLQCPNLTESRAC
DAPISCPLGEEEYTFSLKVGPWSKCRLPHLKEINPSGRTVLDFNSDSNERVTFKHQSYKAHHHSKSWAIEIGYQTRQ
VSCTRSDGQNAMLSLCLQDSFPLTVQSCIMPKDCETSQWSSWSPCSKTCRSGSLLPGFRSRSRNVKHMAIGGGKECP
ELLEKEACIVEGELLQQCPRYSWRTSEWKECQVSLLLEQQDPHWHVTGPVCGGGIQTREVYCAQSVPAAAALRAKEV
SRPVEKALCVGPAPLPSQLCNIPCSTDCIVSSWSAWGLCIHENCHEPQGKKGFRTRQRHVLMESTGPAGHCPHLVES
VPCEDPMCYRWLASEGICFPDHGKCGLGHRILKAVCQNDRGEDVSGSLCPVPPPERKSCEIPCRMDCVLSEWTEWS
SCSQSCSNKNSDGKQTRSRTILALAGEGGKPCPPSQALQEHRLCNDHSCMQLHWETSPWGPCSEDTLVTALNATIGW
NGEATCGVGIQTRRVFCVKSHVGQVMTKRCPDSTRPETVRPCFLPCKKDCIVTAFSEWTPCPRMCQAGNATVKQSRY
RIIIQEAANGGQECPDTLYEERECEDVSLCPVYRWKPQKWSPCILVPESVWQGITGSSEACGKGLQTRAVSCISDDN
RSAEMMECLKQTNGMPLLVQECTVPCREDCTFTAWSKFTPCSTNCEATKSRRRQLTGKSRKKEKCQDSDLYPLVETE
LCPCDEFISQPYGNWSDCILPEGRREPHRGLRVQADSKECGEGLRFRAVACSDKNGRPVDPSFCSSSGYIQEKCVIP
CPFDCKLSDWSSWGSCSSSCGIGVRIRSKWLKEKPYNGGRPCPKLDLKNQAQVHEAVPCYSECNQYSWVVEHWSSCK
INNELRSLRCGGGTQSRKIRCVNTADGEGGAVDSNLCNQDEIPPETQSCSLMCPNECVMSEWGLWSKCPQSCDPHTM
QRRTRHLLRPSLNSRTCAEDSQVQPCLLNENCFQFQYNLTEWSTCQLSENAPCGQGVRTRLLSCVCSDGKPVSMDQC
EQHNLEKPQRMSIPCLVECVVNCQLSGWTAWTECSQTCGHGGRMSRTRFIIMPTQGEGRPCPTELTQEKTCPVTPCY
SWVLGNWSACKLEGGDCGEGVQIRSLSCMVHSGSISHAAGRVEDALCGEMPFQDSILKQLCSVPCPGDCHLTEWSEW
STCELTCIDGRSFETVGRQSRSRTFIIQSFENQDSCPQQVLETRPCTGGKCYHYTWKASLWNNNERTVWCQRSDGVN
VTGGCSPQARPAAIRQCIPACRKPFSYCTQGGVCGCEKGYTEIMRSNGFLDYCMKVPGSEDKKADVKNLSGKNRPVN
SKIHDIFKGWSLQPLDPDGRVKIWVYGVSGGAFLIMIFLIFTSYLVCKKPKPHQSTPPQQKPLTLAYDGDLDMN

Transmembrane domain.

amino acids 1486-1506

N-glycosylation sites.

amino acids 78-81, 118-121, 147-150, 611-614, 685-688, 770-773, 861-864, 1116-
1119, 1238-1241, 1386-1389, 1454-1457

Tyrosine kinase phosphorylation sites.

amino acids 834-841, 953-960

N-myristoylation sites.

amino acids 361-366, 639-644, 746-751, 749-754, 938-943, 1013-1018, 1031-1036,
1196-1201, 1237-1242, 1246-1251, 1384-1389, 1417-1422, 1418-1423, 1421-1426,
1490-1495

Amidation sites.

amino acids 123-126, 433-436, 869-872

Growth factor and cytokines receptors family signature 2.

amino acids 268-274, 1061-1067, 1304-1310

Thrombospondin type 1 domains.

amino acids 111-160, 268-326, 416-470, 533-588, 593-658, 669-723, 803-851, 932-
983, 1061-1110, 1181-1231, 1304-1360

FIGURE 149

EKPVRKQTPPTTQIHCGPPKPVLSPSFKTPATPLGLSTSTGHMLMPLCGLLWWWWCCCSGWYCYGLCAPAPQMLRHQ
GLLKCRCRMLFNDLKVFLLRRPPQAPLPMHGDPQPPGLAANNTLPALGAGGWAGWRGPREVVGREPPPVPPPPPLPP
SSVEDDWGGPATEPPASLLSSASSDDFCKEKTEDRYSLGSSLDSGMRTPLCRICFQGPEQGELLSPCRCDGSVKCTH
QPCLIKWISERGCWSCELCYYKYHVIAISTKNPLQWQAISLTVIEKVQVAAAILGSLFLIASISWLIWSTFSPSARW
QRQDLLFQICYGMYGFMDVVCIGLIIHEGPSVYRIFKRWQAVNQQWKVLNYDKTKDLEDQKAGGRTNPRTSSSTQAN
IPSSEEETAGTPAPEQGPAQAAGHPSGPLSHHHCAYTILHILSHLRPHEQRSPPGSSRELVMRVTTVN

Transmembrane domains.
amino acids 275-295, 316-336

N-glycosylation site.
amino acids 118-121

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 5-8

Tyrosine kinase phosphorylation site.
amino acids 183-190

N-myristoylation sites.
amino acids 35-40, 78-83, 114-119, 127-132, 162-167, 225-230, 243-248

FIGURE 150

QEQGDKMMEEYSLEKNERACIDFAISAKPLTRHMPQNKQSFQYRMWQFVVSPPFEYTIMAMIALNTIVLMMKFYGAS
VAYENALRVFNIVFTSLFSLECVLKVMAFGILNYFRDAWNIFDFVTVLGSITDILVTEFGNNFINLSFLRLFRAARL
IKLLRQGYTIRILLWTFVQSFKALPYVCLLIAMLFFIYAIIGMQVFGNIGIDVEDEDSDEDEFQITEHNNFRTFFQA
LMLLFRSATGEAWHNIMLSCLSGKPCDKNSGILTRECGNEFAYFYFVSFIFLCSFLMLNLFVAVIMDNFEYLTRDSS
ILGPHHLDEYVRVWAEYDPAACGRIHYKDMYSLLRVISPPLGLGKKCPHRVACKRLLRMDLPVADDNTVHFNSTLMA
LIRTALDIKIAKGGADKQQMDAELRKEMMAIWPNLSQKTLDLLVTPHKSTDLTVGKIYAAMMIMEYYRQSKAKKLQA
MREEQDRTPLMFQRMEPPSPTQEGGPGQNALPSTQLDPGGALMAHESGLKESPSWVTQRAQEMFQKTGTWSPEQGPP
TDMPNSQPNSQSVEMREMGRDGYSDSEHYLPMEGQGRAASMPRLPAENQTISDTSPMKRSASVLGPKARRLDDYSLE
RVPPEENQRHHQRRRDRSHRASERSLGRYTDVDTGLGTDLSMTTQSGDLPSKERDQERGRPKDRKHRQHHHHHHHH
HPPPPPDKDRYAQERPDHGRARARDQRWSRSPSEGREHMAHRQGSSSVSGSPAPSTSGTSTPRRGRRQLPQTPSTPRP
HVSYSPVIRKAGGSGPPQQQQQQQQQQQQAVARPGRAATSGPRRYPGPTAEPLAGDRPPTGGHSSGRSPRMERRVPG
PARSESPRACRHGGARWPASGPHVSEGPPGPRHHGYYRGSDYDEADGPGSGGGEEAMAGAYDAPPPVRHASSGATGR
SPRTPRASGPACASPSRHGRRLPNGYYPAHGLARPRGPGSRKGLHEPYSESDDDWCN

Transmembrane domains.
amino acids 49-69, 86-106, 114-134, 131-151, 167-187, 183-203, 222-242, 274-294

N-glycosylation sites.
amino acids 142-145, 380-383, 419-422, 587-590

Glycosaminoglycan attachment site.
amino acids 897-900

N-myristoylation sites.
amino acids 75-80, 269-274, 502-507, 573-578, 653-658, 831-836, 896-901, 899-
904, 963-968

Amidation sites.
amino acids 351-354, 756-759, 942-945

Ion transport protein homology.
amino acids 83-296

FIGURE 151

MRPVALLLLPSLLALLAHGLSLEAPTVGKGQAPGIEETDGELTAAPTPEQPERGVHFVTTAPTLKLLNHHPLLEEFL
HEGLEKGDEELRPALSFQPDPPAPFTPSPLPRLANQDSRPVFTSPTPAMAAVPTQPQSKEGPWSPDPESESPMLRIT
APLPPGPSMAVPTLGPGEIASTTPPSRAWTPTQEGPGDMGRPWVAEVVSQGAGIGIQGTITSSTASGDDEETTTTTT
IITTTITTVQTPGPCSWNFSGPEGSLDSPTDLSSPTDVGLDCFFYISVYPGYGVEIKVKNISLREGETVTVEGLGGP
DPLPLANQSFLLRGQVIRSPTHQAALRFQSLPPPAGPGTFHFHYQAYLLSCHFPRRPAYGDVTVTSLHPGGSARFHC
ATGYQLKGARHLTCLNATQPFWDSKEPVCIGECPGVIRNATTGRIVSPGFPGNYSNNLTCHWLLEAPEGQRLHLHFE
KVSLAEDDDRLIIRNGDNVEAPPVYDSYEVEYLPIEGLLSSGKHFFVELSTDSSGAAAGMALRYEAFQQGHCYEPFV
KYGNFSSSTPTYPVGTTVEFSCDPGYTLEQGSIIIECVDPHDPQWNETEPACRAVCSGEITDSAGVVLSPNWPEPYG
RGQDCIWGVHVEEDKRIMLDIRVLRIGPGDVLTFYDGDDLTARVLGQYSGPRSHFKLFTSMADVTIQFQSDPGTSVL
GYQQGFVIHFFEVPRNDTCPELPEIPNGWKSPSQPELVHGTVVTYQCYPGYQVVGSSVLMCQWDLTWSEDLPSCQRV
TSCHDPGDVEHSRRLISSPKFPVGATVQYICDQGFVLMGSSILTCHDRQAGSPKWSDRAPKCLLEQLKPCHGLSAPE
NGARSPEKQLHPAGATIHFSCAPGYVLKGQASIKCVPGHPSHWSDPPPICRAASLDGSTTVAAWMVAKAPAASSTLD
AAHIAAAIFLPLVAMVLLVGGVYFYFSRLQGKSSLQLPRPRPRPYNRITIESAFDNPTYETGSLSFAGDERI

Signal sequence.
amino acids 1-17

Transmembrane domain.
amino acids 928-948

N-glycosylation sites.
amino acids 249-252, 291-294, 315-318, 401-404, 424-427, 438-441, 442-445, 543-
546, 585-588, 709-712

N-myristoylation sites.
amino acids 30-35, 34-39, 80-85, 205-210, 209-214, 212-217, 420-425, 437-442,
499-504, 517-522, 542-547, 733-738

Sushi domain (SCR repeat).
amino acids 359-414, 534-591, 712-767, 773-832, 840-897

CUB domains.
amino acids 418-526, 595-703

FIGURE 152

MVQKSRNGGVYPGPSGEKKLKVGFVGLDPGAPDSTRDGALLIAGSEAPKRGSILSKPRAGGAGAGKPPKRNAFYRKL
QNFLYNVLERPRGWAFIYHAYVFLLVFSCLVLSVFSTIKEYEKSSEGALYILEIVTIVVFGVEYFVRIWAAGCCCRY
RGWRGRLKFARKPFCVIDIMVLIASIAVLAAGSQGNVFATSALRSLRFLQILRMIRMDRRGGTWKLLGSVVYAHSKE
LVTAWYIGFLCLILASFLVYLAEKGENDHFDTYADALWWGLITLTTIGYGDKYPQTWNGRLLAATFTLIGVSFFALP
AGILGSGFALKVQEQHRQKHFEKRRNPAAGLIQSAWRFYATNLSRTDLHSTWQYYERTVTVPMYRLIPPLNQLELLR
NLKSKSGLAFRKDPPPEPSPSQKVSLKDRVFSSPRGVAAKGKGSPQAQTVRRSPSADQSLEDSPSKVPKSWSFGDRS
RARQAFRIKGAASRQNSEEASLPGEDIVDDKSCPCEFVTEDLTPGLKVSIRAVCVMRFLVSKRKFKESLRPYDVMDV
IEQYSAGHLDMLSRIKSLQSRVDQIVGRGPAITDKDRTKGPAEAELPEDPSMMGRLGKVEKQVLSMEKKLDFLVNIY
MQRMGIPPTETEAYFGAKEPEPAPPYHSPEDSREHVDRHGCIVKIVRSSSSTGQKNFSAPPAAPPVQCPPSTSWQPQ
SHPRQGHGTSPVGDHGSLVRIPPPPAHERSLSAYGGGNRASMEFLRQEDTPGCRPPEGTLRDSDTSISIPSVDHEEL
ERSFSGFSISQSKENLDALNSCYAAVAPCAKVRPYIAEGESDTDSDLCTPCGPPPRSATGEGPFGDVGWAGPRK

Transmembrane domains.
amino acids 89-109, 127-147, 165-185, 184-204, 229-249, 295-315

N-glycosylation sites.
amino acids 350-353, 672-675

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 49-52

Tyrosine kinase phosphorylation site.
amino acids 120-127

N-myristoylation sites.
amino acids 26-31, 30-35, 51-56, 60-65, 61-66, 186-191, 189-194, 310-315, 313-318, 338-343, 428-433, 507-512, 621-626, 729-734, 730-735

Ion transport protein homology.
amino acids 125-312

KCNQ voltage-gated potassium channel homology.
amino acids 417-653

FIGURE 153

MVFAHRMDNSKPHLIIPTLLVPLQNRSCTETATPLPSQYLMELSEEHSWMSNQTDLHYVLKPGEVATASIFFGILWL
FSIFGNSLVCLVIHRSRRTQSTTNYFVVSMACADLLISVASTPFVLLQFTTGRWTLGSATCKVVRYFQYLTPGVQIY
VLLSICIDRFYTIVYPLSFKVSREKAKKMIAASWIFDAGFVTPVLFFYGSNWDSHCNYFLPSSWEGTAYTVIHFLVG
FVIPSVLIILFYQKVIKYIWRIGTDGRTVRRTMNIVPRTKVKTIKMFLILNLLFLLSWLPFHVAQLWHPHEQDYKKS
SLVFTAITWISFSSSASKPTLYSIYNANFRRGMKETFCMSSMKCYRSNAYTITTSSRMAKKNYVGISEIPSMAKTIT
KDSIYDSFDREAKEKKLAWPINSNPPNTFV

Transmembrane domains.
amino acids 67-87, 105-125, 147-167, 184-204, 225-245, 273-293, 311-331

N-glycosylation sites.
amino acids 25-28, 52-55

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 306-309

N-myristoylation sites.
amino acids 134-139, 220-225, 340-345

7 transmembrane receptor (rhodopsin family) homology.
amino acids 82-330

FIGURE 154

MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKC
NGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQE
PLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRR
SASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSDLCVHTASGPEDVALYVG
LIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQ
DGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGIS
LLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQS
CEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLH
DTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT
LERVSPSTSDLACKLWVWQVEGDGQSFS1NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPP
CRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC

Signal sequence.
amino acids 1-25
Transmembrane domain.
amino acids 304-324
N-glycosylation sites.
amino acids 107-110, 218-221, 287-290, 441-444, 682-685, 725-728, 816-819
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 229-232, 536-539
Tyrosine kinase phosphorylation site.
amino acids 212-219
N-myristoylation sites.
amino acids 6-11, 13-18, 24-29, 133-138, 175-180, 255-260, 274-279, 405-410,
509-514, 717-722, 828-833
Growth factor and cytokines receptors family signature 2.
amino acids 243-249, 246-252
ZU5 domain.
amino acids 439-542
Death domain.
amino acids 761-841

FIGURE 155

MAQALLVPPGPGPESFRLFTRESLAAIEKRAAEEKAKKPKKEQDNDDENKPKPNSDLEAGKNLPFIYGDIPPEMVSEPL
EDLDPYYINKKTFIVMNKGKAIFRFSATSALYILTPLNPVRKIAIKILVHSLFSMLIMCTILTNCVFMTLSNPPDWT
KNVEYTFTGIYTFESLIKILVRGFCLEDFTFLRDPWNWLDFSVIVMAYVTEFVSLGNVSALRTFRVLRALKTISVIP
GLKTIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCLQWPPSDSAFETNTTSYFNGTMDSNGTFVNV
TMSTFNWKDNIGDDSHFYVLDGQKDPLLCGNGSDAGQCPEGYICVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDYW
ENLYQLTLRAAGKTYTIFFVLVIFLGSFYLVNLILAVVAMAYEEQNQATLEEAEQKEAEFQQMLEQLKKQQEEAQAV
AAASAASRDFSGIGGLGELLESSSEASKLSSKSAKEWRNRRKKRRQREHLEGNNKGERDSFPKSESEDSVKRSSFLF
SMDGNRLTSDKKFCSPHQSLLSIRGSLFSPRRNSKTSIFSFRGRAKDVGSENDFADDEHSTFEDSETRRDSLFEPHR
HGERRNSNGTTTETEVRKRRLSSYQISMEMLEDSSGRQRAVSIASILTNTMEELEESRQKCPPCWYRFANVFLIWDC
CDAWLKVKHLVNLIVMDPFVDLAITICIVLNTLFMAMEHYPMTEQFSSVLTVGNLVFTGIFTAEMVLKIIAMDPYYY
FQEGWNIFDGIIVSLSLMELGLSNVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIF
AVVGMQLFGKSYKECVCKINDDCTLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQTMCLIVFMLVMVIGNL
VVLNLFLALLLSSFSSDNLAATDDDNEMNNLQIAVGRMQKGIDYVKNKMRECFQKAFFRKPKVIEIHEGNKIDSCMS
NNTGIEISKALNYLRDGNGTTSGVGTGSSVEKYVIDENDYMSFINNPSLTVTVPIAVGESDFENLNTEEFSSESELE
ESKEKLNATSSSEGSTVDVVLPREGEQAETEPEEDLKPEACFTEGCIKKFPFCQVSTEEGKGKIWWNLRKTCYSIVE
HNWFETFIVFMILLSSGALAFEDIYIEQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGFQTYFTNAWCRLDFLIVD
VSLVSLVANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVNALVGAIPSIMNVLLVCLIFWLIFSIMGVNLFA
GKFYHCVNMTTGNMFDISDVNNLSDCQALGKQARWKNVKVNFDNVGAGYLALLQVVS

Transmembrane domains.

amino acids 124-144, 249-269, 402-422, 707-727, 740-760, 778-798, 821-841,
837-857, 874-894, 914-934, 1155-1175, 1192-1212, 1224-1244, 1240-1260,
1269-1289, 1285-1305

N-glycosylation sites.

amino acids 211-214, 290-293, 296-299, 302-305, 307-310, 339-342, 624-627,
835-838, 1002-1005, 1019-1022, 1085-1088, 1317-1320, 1331-1334

Glycosaminoglycan attachment sites.

amino acids 473-476, 1023-1026

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 245-248, 533-536, 570-573, 607-610, 620-623, 635-638

Tyrosine kinase phosphorylation site.

amino acid 1033-1041

N-myristoylation sites.

amino acids 210-215, 271-276, 297-302, 303-308, 340-345, 514-519, 543-548,
752-757, 780-785, 827-832, 831-836, 907-912, 1005-1010, 1018-1023, 1020-1025,
1024-1029, 1026-1031, 1249-1254, 1281-1286

ATP/GTP-binding site motif A (P-loop).

amino acids 851-858

Ion transport protein.

amino acids 153-423, 742-934, 1190-1365

FIGURE 156

MAQALLVPPGPESFRLFTRESLAAIEKRAAEEKAKKPKKEQDNDDENKPKPNSDLEAGKNLPFIYGDIPPEMVSEPL
EDLDPYYINKKTFIVMNKGKAIFRFSATSALYILTPLNPVRKIAIKILVHSLFSMLIMCTILTNCVFMTLSNPPDWT
KNVEYTFTGIYTFESLIKILARGFCLEDFTFLRDPWNWLDFSVIVMAYVTEFVDLGNVSALRTFRVLRALKTISVIP
GLKTIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCLQWPPSDSAFETNTTSYFNGTMDSNGTFVNV
TMSTFNWKDYIGDDSHFYVLDGQKDPLLCGNGSDAGQCPEGYICVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDYW
ENLYQLTLRAAGKTYMIFFVLVIFLGSFYLVNLILAVVAMAYEEQNQATLEEAEQKEAEFQQMLEQLKKQQEEAQAV
AAASAASRDFSGVGGLGELLESSSEASKLSSKGAKEWRNRRKKRRQREHLEGNNKGERDSFPKSESEDSVKRSSFLF
SMDGNRLTSDKKFCSPHQSLLSIRGSLFSPRRNSKTSIFSFRGRAKDVGSENDFADDEHSTFEDSESRRDSLFVPHR
HGERRNSNGTTTETEVRKRRLSSYQISMEMLEDSSGRQRAVSIASILTNTMEELEESRQKCPPCWYRFANVFLIWDC
CDAWLKVKHLVNLIVMDPFVDLAITICIVLNTLFMAMEHYPMTEQFSSVLTVGNLVFTGIFTAEMVLKIIAMDPYYY
FQEGWNIFDGIIVSLSLMELGLSNVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIF
AVVGMQLFGKSYKECVCKINDDCTLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQTMCLIVFMLVMVIGNL
VVLNLFLALLLSSFSSDNLAATDDDNEMNNLQIAVGRMQKGIDYVKNKMRECFQKAFFRKPKVIEIHEGNKIDSCMS
NNTGIEISKELNYLRDGNGTTSGVGTGSSVEKYVIDENDYMSFINNPSLTVTVPIAVGESDFENLNTEEFSSESELE
ESKEKLNATSSSEGSTDVVLPREGEQAETEPEEDFKPEACFTEGCIKKFPFCQVSTEEGKGKIWWNLRKTCYSIVE
HNWFETFIVFMILLSSGALAFEDIYIEQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGFQTYFTNAWCWLDFLIVD
VSLVSLVANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVNALVGAIPSIMNVLLVCLIFWLIFSIMGVNLFA
GKFYHCVNMTTGNMFDISDVNNLSDCQALGKQARWKNVKVNFDNVGAGYLALLQVATFKGWMDIMYAAVDSRDVKLQ
PVYEENLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFGGQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPA
NKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQGKYMTLVLSRINLVFIVLFTGEFVLKLVSLRHYYFTIGW
NIFDFVVVILSIVGMFLAEMIEKYSVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMSLPALFNIGLLLFLVM
FIYAIFGMSNFAYVKKEAGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPILNSAPPDCDPDTIHPGSSVKGDRGD
PSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEKFDPDATQFIEFSKLSDFAAAL
DPPLLIAKPNKVQLIAMDLPMVSGDRIHCLDILFAFTKRVLCESGEMDALRIQMEDRFMASNPSKVSYEPITTTLKR
KQEEVSAAIIQRNFRCYLLKQRLKNISSNYNKEAIKGRIDLPIKQDMIIDKLNGNSTPEKTDGSSSTTSPPSYDSVT
KPDKEKFEKDKPEKESKGKEVRENQK

Transmembrane domains.
amino acids 124-144, 249-269, 402-422, 707-727, 740-760, 778-798, 821-841, 837-
857, 874-894, 914-934, 1155-1175, 1191-1211, 1224-1244, 1240-1260, 1269-1289,
1285-1305, 1351-1371, 1396-1416, 1472-1492, 1507-1527, 1538-1558, 1556-1576,
1590-1610, 1606-1626, 1648-1668, 1700-1720
N-glycosylation sites.
amino acids 211-214, 290-293, 296-299, 302-305, 307-310, 339-342, 624-627, 835-
838, 1002-1105, 1019-1022, 1085-1088, 1317-1320, 1331-1334, 1724-1727, 1873-1876
Glycosaminoglycan attachment sites.
amino acids 473-476, 1023-1026
Tyrosine kinase phosphorylation sites.
amino acids 1033-1041, 1368-1375
N-myristoylation sites.
amino acids 210-215, 271-276, 297-302, 303-308, 340-345, 514-519, 543-548, 752-
757, 780-785, 827-832, 831-836, 907-912, 1005-1010, 1018-1023, 1020-1025, 1024-
1029, 1026-1031, 1249-1254, 1281-1286, 1406-1411, 1554-1559, 1911-1916
Cell attachment sequence.
amino acids 1692-1694
ATP/GTP-binding site motif A (P-loop).
amino acids 851-858
Ion transport protein homology
amino acids 153-423, 742-934, 1190-1418, 1508-1721
Calmodulin-binding motif.
amino acids 1852-1872
ATP synthase homology.
amino acids 399-565
ABC-2 type transporter homology.
amino acids 1421-1636
Alg9-like mannosyltransferase family homology.
amino acids 1223-1667

FIGURE 157

MGAAAAQAPLGLPAASARLLLLATSVLLLFAFSLPGSRASNQPPGGGGGTGGDCPGGKGKSINCSELNVRESDVRVC
DESSCKYGGVCKEDGDGLKCACQFQCHTNYIPVCGSNGDTYQNECFLRRAACKHQKEITVIARGPCYSDNGSGSGEG
EEEGSGAEVHRKHSKCGPCKYKAECDEDAENVGCVCNIDCSGYSFNPVCASDGSSYNNPCFVREASCIKQEQIDIRH
LGHCTDTDDTSLLGKKDDGLQYRPDVKDASDQREDVYIGNHMPCPENLNGYCIHGKCEFIYLLRRASCRCESGYTGQ
HCEKTDFSILYVVPSRQKLTHVLIAAIIGAVQIAIIVAIVMCITRKCPKNNRGRRQKQNLGHFTSDTSSRMV

Signal sequence.
amino acids 1-39
Transmembrane domain.
amino acids 329-349
N-glycosylation sites.
amino acids 63-66, 147-150
Glycosaminoglycan attachment sites.
amino acids 149-152, 151-154
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 165-168, 295-298
Tyrosine kinase phosphorylation site.
amino acids 246-253
N-myristoylation sites.
amino acids 2-7, 11-16, 36-41, 45-50, 46-51, 47-52, 48-53, 94-99, 148-153,
150-155, 187-192, 207-212
Amidation sites.
amino acids 244-247, 360-363
EGF-like domain cysteine pattern signature.
amino acids 299-310
Kazal-type serine protease inhibitor domain.
amino acids 99-143, 190-235

FIGURE 158

```
MLPEQLYFLQSPPEEEPEYHPDASAQELNVRESDVRVCDESSCKYGGVCKEDGDGLKCACQFQCHTNYIPVCGSNGD
TYQNECFLRRAACKHQKEITVIARGPCYSDNGSGSGEGEEEGSGAEVHRKHSKCGPCKYKAECDEDAENVGCVCNID
CSGYSFNPVCASDGSSYNNPCFVREASCIKQEQIDIRHLGHCTDTDDTSLLGKKDDGLQYRPDVKDASDQREDVYIG
NHMPCPENLNGYCIHGKCEFIYSTQKASCRCESGYTGQHCEKTDFSILYVVPSRQKLTHVLIAAIIGAVQIAIIVAI
VMCITRKCPKNNRGRRQKQNLGHFTSDTSSRMV
```

Transmembrane domain.
amino acids 290-310

N-glycosylation site.
amino acids 108-111

Glycosaminoglycan attachment site.
amino acids 110-113, 112-115

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 126-129

Tyrosine kinase phosphorylation site.
amino acids 207-214

N-myristoylation sites.
amino acids 55-60, 109-114, 111-116, 148-153, 168-173

Amidation sites.
amino acids 205-208, 321-324

EGF-like domain cysteine pattern signature.
amino acids 260-271

Kazal-type serine protease inhibitor domains.
amino acids 60-104, 151-196

EGF-like domains.
amino acids 38-64, 236-271